
IN DEO NON ARMIS FIDO

3-12-19

To Dr J L Morse
with Compliments
of the Publishers
L. A. D.

PHYSIOLOGY AND BIOCHEMISTRY IN MODERN MEDICINE

BY

J. J. R. MACLEOD, M.B.

PROFESSOR OF PHYSIOLOGY IN THE UNIVERSITY OF TORONTO, TORONTO, CANADA; FORMERLY PROFESSOR OF PHYSIOLOGY IN THE WESTERN RESERVE UNIVERSITY, CLEVELAND, OHIO

ASSISTED BY ROY G. PEARCE, B.A., M.D.

Director of the Cardiorespiratory Laboratory of Lakeside Hospital, Cleveland, Ohio

AND BY OTHERS

WITH 233 ILLUSTRATIONS, INCLUDING 11 PLATES IN COLORS

ST. LOUIS

C. V. MOSBY COMPANY

1918

Press of
C. V. Mosby Company
St. Louis

TO

M. W. M.

PREFACE

The necessity of allotting the various subjects of the medical curriculum to different periods, so that the more strictly scientific subjects are completed in the earlier years, has the great disadvantage that the student, being no longer in touch with laboratory work, fails to employ the scientific knowledge with full advantage in the solution of his clinical problems. He is apt to regard the first two or three years in the laboratory departments as inconsequential in comparison with the supposedly more practical instruction offered during the subsequent clinical years. He is taught by his laboratory instructors to observe accurately, and to correlate the observed facts, so that he may be enabled to draw conclusions as to the manner of working of the various functions of the animal body in health, and before proceeding to his clinical studies, he is required to show a proficiency in scientific knowledge, because it is recognized that this must serve as the basis upon which his knowledge of disease is to be built. When the clinic is reached, however, the methods of the scientist are not infrequently cast aside and an understanding of disease is sought for largely by the empirical method; namely, by the endeavor to see and examine innumerable patients, to diagnose the case according to the grouping of the signs and symptoms, and to treat it by the prescribed methods of experience. So much has to be learned and so much has to be seen during the clinical years, that the student gives little thought to the nature of the functional disturbance which is responsible for the symptoms; he fails to realize that after all, there is no essential difference between the condition brought about in his patient by some pathologic lesion, and that which may be produced in the laboratory by experimental procedures, by drugs or by toxins. It must of course be recognized that just as the science of medicine originated by the grouping of symptoms into more or less characteristic diseases for which the most favorable method of treatment had to be discovered by experience, so must a certain part of the medical training be more or less empirical but it should at the same time be realized that such a method is only a means to an end, and that the real understanding of disease can be acquired only when every abnormal condition is interpreted as a primary or secondary consequence of some perverted bodily function, and when the training in observation and the inductive method is carried from the laboratory into the clinic.

It is a constant experience of clinical instructors who would employ scientific methods of instruction, that they find the students not only indifferent to an analysis of their cases from the functional standpoint, but also that they are too inadequately prepared in fundamental physiologic knowledge, to make the analysis possible. The student may have a superficial acquaintance with the main facts of physiologic science but have failed to acquire the enquiring habit of mind which will enable him, through reflection, comparison, and personal research, to apply the knowledge in practical medicine and surgery.

For this lack of correlation between the laboratory and clinical studies, the clinical instructors are not alone responsible. The laboratory courses are frequently given without any attempt being made to show the student the bearing of the subject in the interpretation of disease, or to train him so that in his later years he may be able to adapt the methods of investigation which he learned in the laboratory, to the study of morbid conditions. It is self-evident that (without any knowledge of disease) the extent to which the student in the earlier years of the course could be expected to appreciate the clinical significance of what he learns in the laboratory is limited, but this should not deter the instructor from indicating whenever he can, the general application of scientific knowledge in the interpretation of diseased conditions. But the chief remedy of the evil undoubtedly lies partly in the continuance of certain of the laboratory courses into the clinical years, and partly in the study of medical literature in which the application of physiology and biochemistry in the practice of medicine is emphasized.

Notwithstanding the sufficient number of excellent textbooks in physiology available to the medical student, there is none in which particular emphasis is laid upon the application of the subject in the routine practice of medicine. In the present volume the attempt is made to meet such a want, by reviewing those portions of physiology and biochemistry which experience has shown to be of especial value to the clinical investigator. The work is not intended to be a substitute, either for the regular textbooks in physiology, or for those in functional pathology. It is supplementary to such volumes. It does not start like the modern test in functional pathology, with a consideration of the diseased condition, and then proceed to analyze the possible causes and consequences of the disturbances of function which this exhibits; but it deals with the present-day knowledge of human physiology in so far as this can be used in a general way to advance the understanding of disease. In a sense it is therefore an advanced text in physiology for those about to enter upon their clinical instruction, and at the same

time, a review for those of a maturer clinical experience who may desire to seek the physiological interpretation of diseased conditions.

In attempting to fulfil these requirements, it has been deemed essential to go back to the fundamentals of the subject, and to explain as simply as possible the physical and physicochemical principles upon which so large a part of physiological knowledge depends. Physiology may be considered as an application of the known laws and facts of physics and chemistry to explain the functions of living matter, and it is only after the extent to which this application can be made has been appreciated, that the knowledge may be used to serve as the foundation upon which a superstructure of clinical knowledge can be built.

In order that the volume might be maintained of reasonable size, it has been necessary to select certain parts of the subject for particular emphasis, the basis of selection being the degree to which our knowledge clearly shows the value of the application of physiological methods both of observation and of thought in the study of diseased conditions. This has not been done to the extent of omitting the apparently less essential parts, for these have been treated in sufficient detail to link the others together so as to preserve a logical continuity, and show the bearing of one field of knowledge on another. There are however certain parts of the science, particularly the physiology of nerve and muscle, of the special senses, and of reproduction, for which application in the general fields of medicine and surgery is limited, and these parts have been omitted entirely. It has been judged that this perhaps somewhat arbitrary selection is justified on the ground that the ordinary text in physiology covers these subjects sufficiently, except for the specialist, for whom on the other hand, no adequate review would have been possible within the limits of such a volume as this. With reference to biochemistry, no attempt is made to review the properties or describe the characteristic tests of the various chemical ingredients of the body tissues and fluids. This is already sufficiently done in the textbooks on biochemistry, and in the numerous manuals on clinical methods. Biochemical knowledge is treated rather from the physiologist's standpoint, as an integral part of his subject, particular attention, nevertheless, being paid to the far-reaching applications of this latest department of medical science, in the elucidation of many obscure problems of clinical medicine, such as those of diabetes, nephritis, acidosis, goiter and myxedema. To make the volume of value to those who may not have had time or opportunity to familiarize themselves with the technical methods of the physiologist and biochemist as used in the modern clinic, a certain amount of space is devoted to a brief description of the methods that appear at present to be receiving most attention, and to be of greatest value.

Finally, it should be mentioned that the principles of serum diagnosis and therapy are omitted, since these belong to a highly specialized science requiring an intensive training of its own.

In the hope that the volume may be instrumental in arousing sufficient interest to stimulate a more intensive study of the various subjects which it introduces, a brief bibliography is given at the end of each section. The references selected are to papers that are more particularly known to the author; they are not necessarily the most important publications on the subject, but are often chosen because of the useful reviews of previous work contained in them, rather than because of their own originality. Some of the papers, however, are referred to as authority for statements of fact which may arouse in the reader a desire to ponder for himself the evidence upon which these are based. The references are usually divided into two groups, "monographs" and "original papers," and it is only occasionally that specific reference is made to the former in the context. The original papers, on the other hand, are referred to by numbers. With the general field of the subject so well covered by such excellent textbooks as Bayliss' "Principles of General Physiology," Stewart's, Howell's, Starling's, and Halliburton's "Human Physiologies," and Leonard Hill's "Recent and Further Advances in Physiology," the author has felt free to pick and choose from the monographs and original papers, topics that are ordinarily passed over cursorily in the textbook, and when this has been done, the references are somewhat more extensive. Such is the case for example in the chapters relating to the chemistry of respiration, to the metabolism of carbohydrates and fats, to the problems of dietetics and growth, to the physicochemical basis of neutrality regulation in the animal body, and to the action of enzymes.

Acknowledgment is gratefully made for the assistance and advice in the preparation of the book, particularly to Doctor R. G. Pearce, for the contribution of several chapters, to which his name is attached, and for which he is entirely responsible; and to Doctor E. P. Carter, whose criticisms, after patient perusal of the unfinished manuscript, were of inestimable value in its final revision. Acknowledgment is also made to Doctor R. W. Scott and Professor F. E. Lloyd, for valuable criticism and advice, and to the former for a chapter on the "Clinical Application of Electrocardiographs." To Miss Achsa Parker, M.A., the author owes a great debt of gratitude for the thorough and painstaking way in which she prepared the manuscript for the press, and for her nevertiring endeavors to have the spelling and punctuation in conformity with Webster's Dictionary. For assistance in the preparation of the index thanks are due to Miss Marian Armour and Mrs. MacFarlane,

and for permission to use certain of the figures and illustrations, to the various authors and publishers who granted it. For the excellent management and careful execution of the presswork, the author wishes to thank the publishers, whose courteous and friendly dealings have always made the work easier.

J. J. R. MACLEOD.

University of Toronto,
Toronto, Canada.

CONTENTS

PART I

THE PHYSICOCHEMICAL BASIS OF PHYSIOLOGICAL PROCESSES

PART III

CIRCULATION OF THE BLOOD

CHAPTER XXIII

CHAPTER XXIV

CHAPTER XXV

CHAPTER XXVI

CHAPTER XXVII

CHAPTER XXVIII

CHAPTER XXIX

CHAPTER XXX

PART V

DIGESTION

CHAPTER LVI

CHAPTER LVII

PART VI

THE EXCRETION OF URINE

CHAPTER LVIII

CHAPTER LIX

PART VII

METABOLISM

CHAPTER LX

CHAPTER LXI

CHAPTER LXII

PART VIII

THE ENDOCRINE ORGANS, OR DUCTLESS GLANDS

CHAPTER LXXXI

PAGE

CHAPTER LXXXII

CHAPTER LXXXIII

CHAPTER LXXXIV

CHAPTER LXXXV

PART IX

THE CENTRAL NERVOUS SYSTEM

CHAPTER LXXXVI

CHAPTER LXXXVII

CHAPTER LXXXVIII

CHAPTER C

CHAPTER CI

ILLUSTRATIONS

PHYSIOLOGY AND BIOCHEMISTRY IN MODERN MEDICINE

PART I

THE PHYSICOCHEMICAL BASIS OF PHYSIOLOGICAL PROCESSES

CHAPTER I

GENERAL CONSIDERATIONS

The work of the physiologist consists, in large part, in ascertaining to what extent the known laws of physics and chemistry find application in explaining the phenomena of life. He gathers from the vast storehouse of physical and chemical knowledge whatever is of value in the interpretation of the various mechanisms that work together to compose the living machine, and having added to this knowledge he passes it on for use by those who are concerned in the study and treatment of disease.

Many of the most important steps in the advance of physiologic knowledge in recent years have depended upon the discovery of some hitherto unknown physical or chemical law, or upon the elaboration of some accurate method for the measurement of the phenomena upon which these or previously known laws depend. The discoveries of van 't Hoff, Arrhenius, and Ostwald of the so-called laws of solution were soon followed by important observations on their relationship to the movement of fluids and dissolved substances through cell membranes; the discoveries of Hardy, Willard Gibbs, etc., of the behavior of colloids and of the phenomena of surface tension found application in explaining many hitherto inexplicable peculiarities in the activities of ferments; the discovery by Nernst, etc., of methods for the measurement of the electro-motive force of dissolved substances was applied to determine the actual reaction or hydrogen-ion concentration of animal

fluids, and to explain the generation of the electric currents which accompany muscular, nervous, and glandular activity.

It would be out of place here to devote much space to a detailed account of such matters. They belong more properly in the domain of general than in that of human physiology. General physiology is concerned with the study of the essential nature of the vital processes; whereas human physiology is merely a branch of the subject in which special attention is devoted to the application of the truths of general physiology to the working of the human machine. For the physician and surgeon a knowledge of human physiology is as essential as is a knowledge of the construction of a piece of machinery for the engineer who attempts its repair, but obviously to acquire this knowledge the fundamental principles of general physiology must first of all be understood. For these reasons the introductory chapters are devoted to a brief review of the most important of the physicochemical principles upon which the working of the cell depends.

From the viewpoint of the physical chemist the cell consists of an envelope of more or less permeable material inclosing a dilute solution of crystalline substances in which colloid matter is suspended. It contains, in other words, a solution of crystalloids and colloids, in which these are in a state of equilibrium with each other. This equilibrium is readily altered by various influences that may act on the cell, and the resulting changes manifest themselves outwardly by alterations in the shape and volume of the cell—*growth and motion;* by the extrusion of some of its contents—*secretion;* or by the propagation to other parts of the cell, or its processes, of the state of disturbed equilibrium—*nervous impulse.* Besides the activities that are dependent upon physicochemical changes, purely chemical processes go on in the cell. Many of these consist in the breakdown and oxidation of complex unstable organic molecules, a process identical with that occurring in combustion outside the cell. Others involve the building up, stage by stage, of complex substances out of the elements or out of simpler molecules. Chemical transformations occur in the cell which, in the chemical laboratory, require the most powerful reagents and physicochemical forces, either the strongest of acids, alkalies, oxidizing agents, etc., or extreme degrees of heat, electrical energy, etc. But this is not all, for in the cell these chemical transformations are capable of being guided to a very remarkable degree of nicety so as to produce intermediate products that are used for some special purpose either by the cell that produced them or, after transportation by the blood, etc., by cells in other parts of the organism.

It is customary to speak of the cell as a chemical laboratory, but it

is more than this; it is a laboratory furnished not only with the equipment of the chemist but directed in the harmonious operation of its many activities by a guiding hand which far surpasses anything known to man. Chemical transformations that require for their accomplishment the greatest skill proceed without apparent difficulty in the cell. To what are these changes due? What is the nature of the chemical reagents and forces, and what is the directive influence that guides them in their varied activities? To these, which are among the great questions of general physiology, the reply may be given that the reagents are the ferments or enzymes, and that the directive influence operates through the susceptibility of enzymic activities to changes in the environment in which the enzymes are acting. In many cases these changes can be explained on a physicochemical basis as dependent upon the known laws of mass action or surface tension; in other cases they depend on purely chemical changes in the cell contents, such as changes in reaction or the accumulation of chemical substances that act like poisons on the enzyme. But there are still others that appear to depend on influences which as yet are quite unknown to the physical chemist, such as the changes in cell activity that can be brought about by the nerve impulse.

These preliminary remarks will serve to indicate the problems with which we must first occupy our attention. They concern the physicochemical nature of saline solutions and of colloids, and the general nature of enzyme action. The knowledge which we acquire will be found to be of value, not only because it will help us to understand the nature of the workings of the normal healthy cell, but because, here and there, it will indicate possible causes for derangement in cellular function and suggest rational means by which we may attempt to rectify the fault.

THE PHYSICOCHEMICAL LAWS OF SOLUTION

The Gas Laws

Three fundamental principles of general chemistry serve as the basis for an understanding of the nature of solutions. The first is that if we take a quantity of any gas equal to its molecular weight in grams (called a gram-molecule or for sake of brevity a mol), it will occupy exactly 22.4 liters at standard temperature and pressure; the second is that, as we compress a gas, its pressure will increase in exactly the same proportion as the volume diminishes (the volume of a gas is inversely proportional to its pressure); the third is that all gases expand by 1/273

part of their volume at 0° C. for every degree C. that their temperature is raised.*

The pressure of a gas is measured by connecting a pressure gauge or manometer with the vessel which contains the gas. Now, it is plain that if the 22.4 liters, which is the volume occupied by a gram-molecular quantity, were compressed so as to occupy a volume of 1 liter, its pressure would be 22.4 times that of 1 atmosphere, or 22.4×760 mm. Hg—the temperature remaining constant. Under these conditions we must imagine that the molecules of gas are crowded together by the compression, and if we further conceive of these molecules as being in constant motion, then we can understand why the pressure should increase just in proportion as we confine the space in which they can move.

One other property of gases must be borne in mind—namely, their tendency to diffuse from places where the pressure is high to places where it is low until the pressure is the same throughout.

OSMOTIC PRESSURE

These fundamental facts regarding the behavior of gases suggested to van't Hoff the hypothesis that *molecules of dissolved substances must behave in a similar manner to those* of gases. To put this hypothesis to the test, it is necessary that we have some method for measuring the pressure of dissolved molecules. We can not, as in the case of a gas, use an ordinary manometer, for this would measure only the pressure of the solvent on the walls of its container and would tell us nothing of the pressure of the dissolved molecules. We must use some filter or membrane that will allow the molecules of the solvent but not those of the dissolved substance to pass through it. It is evident that if such a filter is placed, for example, between a solution of sugar in water and water alone, the molecules of the latter will diffuse into the solution until this has become so diluted that the pressure of the dissolved molecules is equal on both sides of the membrane. Such a membrane is called *semipermeable;* the diffusion of molecules through it is called *osmosis,* and the pressure which is generated, *the osmotic pressure.* If we prevent the water molecules from actually diffusing by opposing a pressure which is equal to that with which they tend to diffuse through the membrane, we can tell the magnitude of the osmotic pressure (Fig. 1).

In applying these facts to test the hypothesis that molecules in solution

*This implies that at -273° C. the gas would occupy no volume. Before this temperature is reached, however, the liquefaction of the gas sets in. The temperature -273° C. is known as absolute zero. An observed temperature *plus* 273° is called the absolute temperature. Another way of stating the above law is therefore that the volume is directly proportional to the absolute temperature. At 273° C. the volume of a gas at 0° C. would be doubled, or if expansion were prevented the pressure would be doubled.

obey the same laws as those in gaseous form, we must employ a semipermeable membrane which is rigid enough to withstand the pressure and which forms part of the walls of a closed vessel connected with a manometer. If we place in such an osmometer a solution containing the molecular weight in grams of some substance dissolved in one liter of solvent, a so-called gram-molecular solution, it is obvious that, if the gas laws are to apply, the osmotic pressure should equal that of 22.4 liters of a gas compressed to the volume of one liter; in other words, it should equal 22.4×760 mm. Hg. Although there are very considerable technical difficulties in making a semipermeable membrane that is strong enough to withstand such a pressure, yet this has been accom-

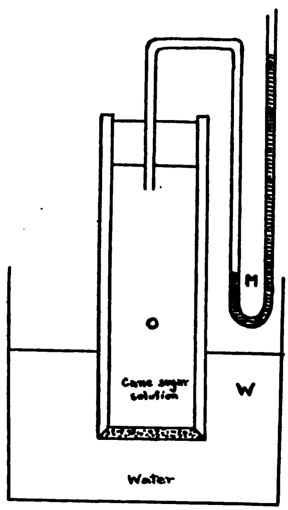


Fig. 1.—Diagram of osmometer. The cylindrical vessel (*O*), with a bottom of unglazed clay, the pores of which are filled with a precipitate of copper ferrocyanide to form a semipermeable membrane, is suspended in an outer vessel, and is closed above by a tightly fitting stopper pierced by a tube leading to a manometer (*M*). *O* contains a strong solution of cane sugar, and *W* contains water. The water molecules tend to pass through the semipermeable membrane into the cane sugar solution, and since the cane sugar molecules can not pass in the opposite direction, the pressure in *O* rises and is recorded in *M*. This equals the osmotic pressure.

plished, and the fundamental principle has therefore been firmly established that substances in solution obey the same laws as gases.

Further proof that the gas laws apply to solutions has been secured by showing that the osmotic pressure (of a dilute solution) is directly proportional to the concentration of the dissolved substance (the solute) and to the absolute temperature. It also obeys the law of partial pressures, which states that the total pressure exerted by a mixture (of gases or dissolved molecules) is the sum of the pressures which each constituent of the mixture would exert were it alone present in the space occupied by the mixture.

Since the osmotic pressure is analogous to the pressure of a gas and is therefore proportional to the molecular concentration (i. e., number of molecules in unit space), it follows that a semipermeable membrane can be used to determine the relative concentration of two solutions of the same substance. When a watery solution of some substance is placed in an osmometer that is surrounded by a similar but more dilute solution, water molecules will diffuse into the osmometer until the pressure is equal on the two sides of the semipermeable membrane; that is, the water will pass from the solution having a lower osmotic pressure into the solution having the higher pressure. When two solutions have the same osmotic pressure, they are said to be *isotonic;* when that of one is greater than that of the other, it is *hypertonic;* and when less, *hypotonic.*

Biological Methods for Measuring Osmotic Pressure

A practical biological application of these principles can very readily be made if, instead of a rigid semipermeable membrane such as that figured in the diagram, we employ one that is extensible and takes the form of a closed sac; then as diffusion of water occurs the sac will either distend when it contains a stronger solution than that outside, or shrivel or crenate when the reverse conditions obtain. Many animal and vegetable protoplasmic membranes are semipermeable, including the envelope of red blood corpuscles. Thus, if we examine blood corpuscles under the microscope and add to them a saline solution of higher osmotic pressure than blood serum, they will visibly diminish in size and become irregular in shape; whereas if the solution is of lower osmotic pressure, they will distend. If no change occurs, the osmotic pressure of the cell contents must equal that of the saline solution in which the cells are immersed, from which it is clear that we can readily determine the magnitude of the osmotic pressure if we know the strength of the saline solution.

Instead of measuring the individual cells under the microscope, we can measure the space they occupy in the fluid in which they are suspended. For this purpose a portion of the suspension is placed in a graduated tube of narrow bore, which is rotated in a horizontal position by a centrifuge after being closed at one end. The graduation at which the upper edge of the column of cells stands after centrifuging is a measure of the relative amount of cells and fluid in the suspension. Having found this value for cells suspended in an isotonic solution, as for blood corpuscles in blood serum, we may then proceed to ascertain it for the same cells suspended in an unknown solution; if we find that the cells occupy a greater volume, the saline solution must have an osmotic pres-

sure that is lower than that of serum in approximate proportion to the readings on the tube in the two cases, and *vice versa.*

The above apparatus, called a hematocrite (Fig. 2) has been very extensively used in the collection of data concerning the relative osmotic pressures of different physiologic fluids.

Hemolysis

Another way for determining the relative osmotic pressure of different solutions consists in placing equal amounts (a few drops) of blood in a series of test tubes containing solutions of different strengths, and after allowing the tubes to stand for some time, noting in which of them laking of the blood corpuscles occurs. In solutions which are isotonic or hypertonic with the contents of the corpuscles, the latter will settle to the bottom of the tube and the supernatant fluid will be untinted with hemoglobin, but in solutions which are distinctly hypotonic, the sediment will be less distinct and the supernatant fluid red.

Fig. 2.—Hematocrite. The graduated glass tubes are filled with the two specimens of blood, or corpuscular suspension, and then rotated rapidly by a centrifuge. The relative heights at which the corpuscular sediment stands in the two tubes is proportional to the osmotic pressures of the fluid in which the corpuscles are suspended.

By noting (1) the lowest concentration (percentage composition) of the solutions in which the corpuscles sink to the bottom and leave the supernatant fluid colorless, and (2) the highest concentration in which the corpuscles when they settle leave the supernatant fluid red, we can determine the limiting concentrations for solutions of different substances. Thus, with bullock's blood the following results were obtained (Hamburger):

Substance	Percentage Strength of Solution in Which: I Supernatant fluid was colorless	II Supernatant fluid was red
KNO_3	1.04	0.96
NaCl	0.60	0.56
K_2SO_4	1.16	1.06
$C_{12}H_{22}O_{11}$ (Cane sugar)	6.29	5.63
CH_3COOH (Pot. acetate)	1.07	1.00
$MgSO_4.7H2O$	3.52	3.26
$CaCl_2$	0.85	0.79

The mean of these limiting concentrations is the *critical concentration* and indicates the strength of each solution that can be added to blood without causing any damage to the corpuscles. This critical concentration is not, as might at first sight be imagined, the same as that which is isotonic with the contents of the corpuscles, but distinctly below it. The reason for this becomes apparent if we observe the behavior of corpuscles suspended in an isotonic solution which is then gradually diluted. As dilution proceeds, the corpuscles distend, until at last their envelopes burst and the hemoglobin is discharged. The limiting concentrations of a given salt vary for different corpuscles; thus, the concentration of sodium chloride solution that just causes laking of frog's blood corpuscles is 0.21 per cent, of human blood 0.47 per cent, and of horse blood 0.68 per cent. It is the strength of the corpuscular envelope rather than variations in the osmotic pressure of the contents that is responsible for these differences.

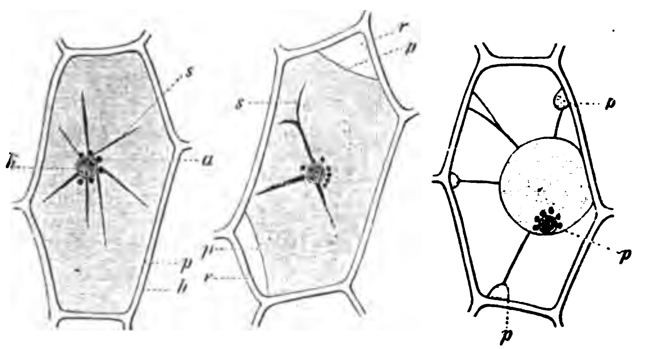
The above described method of hemolysis, as it is called, can not be used for comparisons of osmotic pressure in cases in which the solution contains substances which alter the permeability of the corpuscular envelop; for example, it can not be used when urea, or ammonium salts, or certain toxic bodies are present. This very fact is, however, put to a useful purpose in ascertaining whether a given substance does have a damaging influence on the corpuscular envelope by finding whether hemolysis occurs when we suspend the corpuscles in a solution that *is* isotonic with the corpuscular contents. We can further determine the degree of this toxic influence by estimating by color comparisons (colorimetry) the amount of hemoglobin that has diffused out of the corpuscles.

Plasmolysis

An analogous method for determining osmotic pressure is that of plasmolysis, in which the behavior of certain plant cells is observed microscopically while they are in contact with solutions of different strengths. When the surrounding solution is isotonic with the cell contents, the latter fill the cell and extend up to the more or less rigid cell wall (*A* in Fig. 3); but when the solution is hypotonic, the cell contents become detached from the cell wall at one or more places—plasmolysis (*B* and *C*). The semipermeable membrane in this case is therefore not the cell wall but the layer of protoplasm on the surface of the cell contents. The method can be used only for detecting solutions that are hypertonic, for with those that are hypotonic the cells merely become turgid and exert more pressure on the more or less rigid cell wall. Many of the conclusions that have been drawn from

results obtained by the plasmolytic method have recently been called in question, because no regard has been taken of the power of the colloids of the cell to adsorb (imbibe) water (see page 62).

The methods of hemolysis and plasmolysis have been used for the investigation of many problems in medicine. In the case of certain toxic fluids, such as snake venom, tetanus toxin, etc., determination of the hemolytic power has proved of value in roughly assaying the damaging influence on other cells than blood corpuscles. Studies in hemolysis have also been especially valuable in working out the mechanism by which cellular toxins in general develop their action, and the conditions under which this action may be counteracted, as by the development of



A B C

Fig. 3.—To show plasmolysis in cells from Tradescantia discolor. *A*, normal cell; *B*, plasmolysis in 0.22 M. cane sugar; *C*, pronounced plasmolysis in 1.0 M. KNO_3; *h*, the cell wall; *p*, the protoplasm. (After De Vries.)

antibodies. Furthermore, any solution that is to be injected into the animal body, either intravenously or subcutaneously, should first of all be tested by the above methods in order to find out whether it is isotonic with the body fluids. If a hypertonic solution is injected, it will result in the abstraction of water from the tissue cells, whereas a hypotonic solution will cause the water content of these to increase. Advantage has recently been taken of this water-abstracting effect of hypertonic solutions in the treatment of wounds. By constantly bathing them with strong saline solutions, an outflow of water is set up from the tissue cells that border on the wound, and this tends to bring to the focus of infection the defensive substances that are present in animal fluids.

CHAPTER II

OSMOTIC PRESSURE (Cont'd)

Measurement by Depression of Freezing Point

The limitations in the use of the plasmolytic and hemolytic methods in the precise measurement of the osmotic pressure of the body fluids have rendered it necessary to find some physical method that will be generally applicable. Because of technical difficulties, it is impracticable to measure the pressure directly by employing an osmometer, so that some indirect method, depending on a readily measurable physical property which varies in proportion to the osmotic pressure of the dissolved substances, must be used. Fortunately, one such exists in the property which dissolved substances have in lowering the temperature at which the pure solvent solidifies; the freezing point of pure water, for example, is lowered when substances are dissolved in it, and the extent of this lowering, with certain reservations which will be explained later (page 16), is proportional to the molecular concentration of the solution and independent of the chemical nature of the substance dissolved. This lowering of temperature is designated by the Greek letter Δ, and to measure it a thermometer is used which is not only extremely sensitive but in which the level of the mercury column can be adjusted so that it stands at a convenient level on the scale corresponding to the freezing point of whatever solvent was used in making the solution under investigation (Beckmann's thermometer) (Fig. 4). The exact position on the scale of this thermometer at which the pure solvent freezes having been ascertained, the observation is repeated with the solution whose osmotic pressure is to be determined.

A gram-molecular solution in water (having therefore an osmotic pressure of 170,240 mm. Hg) has a freezing point that is 1.86° C. lower than that of pure water. This is known as the "freezing point constant," and it varies for different solvents, being 3.9 for acetic acid and 4.9 for benzene. If an unknown watery solution is found to have a freezing point that is Δ° C. lower than that of water, its osmotic pressure will equal $\frac{\Delta \times 17.024}{1.86}$ mm. Hg.

The depression of the freezing points produced by the various body fluids has been compared, the objects in view being to see whether osmotic pressure is a property which changes under different physiologic and pathologic conditions, and to find out by comparison of the osmotic pressures of the fluids in contact with a membrane, whether physical forces alone can be held responsible for the transference of substances through it from one fluid to the other.

The Role of Osmosis, Diffusion, and Allied Processes in Physiologic Mechanisms

An account of some of the investigations in which the foregoing methods have been used will illustrate their value in revealing the

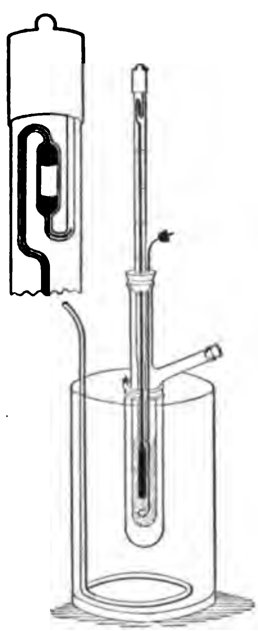


Fig. 4.—Apparatus for measurement of the depression of freezing point of solutions. The solution is placed in the large test tube with the side arm, and in it is suspended the bulb of a Beckmann thermometer with a platinum loop to serve for stirring. The upper end of the mercury column of the thermometer is shown magnified at the upper left corner. The amount of mercury in the thermometer tube can be regulated by tapping the upper end with the thermometer in various positions. The test tube is protected by an outer tube, which is then placed in a vessel containing a freezing mixture.

mechanism involved in the transference of water and dissolved substances through cell membranes, as occurs in absorption of food in the intestine, in the formation of lymph and urine, and so forth. In employing physical methods in the elucidation of such problems, it is always most necessary to proceed with great care, since the physical

chemist works with pure solutions, while the physiologist has to use fluids that are always complicated and frequently very variable in composition. We must simplify the problem as far as possible by having clearly before us the exact nature of the biological problem which a comparison of physicochemical values, such as osmotic pressure, may enable us to elucidate, and we must consider the other physical forces which may assist or modify the particular one we are investigating.

In the physical experiments described above, the semipermeable membrane may be conceived of as composed of pores of such a size that they permit only the smallest of molecules—those of water—to pass through them. Semipermeable membranes with larger pores may, however, exist—that is, membranes which permit water molecules and molecules of simple chemical substances to pass, but hold back those composed of large complex molecules. Such a semipermeable membrane would allow the saline constituents but not the proteins of blood serum to pass. It is, however, no longer semipermeable towards all of the dissolved substances, and the process of *diffusion* through it is more generally designated as one of *dialysis* than of osmosis.

Since the passage of dissolved molecules through membranes depends upon the principle of diffusion, its rate will be proportional to the osmotic pressures of the solutions on the two surfaces of the membrane and to the size of the molecules, small molecules diffusing more quickly than large ones. Suppose a membrane permeable to sodium chloride and water is placed between two fluids containing sodium chloride in solution, but in greater concentration in one of them than in the other: the sodium chloride will diffuse from the stronger to the weaker solution, and water will diffuse still more quickly (because its molecules are smaller) in the opposite direction, until the number of sodium-chloride molecules in a given volume of solution is equal on both sides of the membrane. For a time, therefore, the volume of the stronger solution will increase. The differences which exist in the diffusibility of dissolved molecules are analogous to those which have long been known to exist in the diffusibility of gases, but the relation between rate of diffusibility and molecular weight is not so simple as the ratio between these two quantities in gases. These relationships, however, indicate several further possibilities in the explanation of the mechanism of exchange of substances through membranes, and must not be overlooked, as they often are, in the interpretation of physiologic phenomena. An excellent review of the possible conditions is given by Starling in his "Human Physiology."[4] For example, let us suppose the substances on the two sides of a semipermeable membrane, such as the peritoneal, to be different in diffusibility, as cane sugar,

which does not readily diffuse, and sodium chloride, which diffuses quickly; the osmotic flow will take place from the sodium-chlorid solution to the cane sugar even when the sodium-chloride solution is stronger than the sugar. In such a case, water molecules will pass from the fluid having the higher osmotic pressure (NaCl) toward a fluid in which this is lower (sugar).

Furthermore, the simple laws of osmosis may be upset by an attractive influence of the membrane toward certain substances [due to their becoming dissolved or adsorbed in it (see page 65)] but not toward others. Many membranes of this nature are known to the chemist (e. g., rubber membranes in contact with gases, pyridine solutions, etc.), and it is probable that such a property of selective solubility may play a not unimportant role in the transference of substances across animal membranes (Kahlenberg[5]).

These few conditions which may modify the direction of the osmotic flow, are indicated here to show how involved such problems are, and how careful we must be not to assume that, because a substance is transferred through a living membrane contrary to the simpler laws of osmosia and diffusion, it must involve the expenditure of forces different from those operating in dead membranes.

Another force comes into operation under certain conditions—namely, that of *filtration*. This is a purely mechanical process, in which molecules are forced through the pores of a filter (i. e., membrane) by differences in pressure on its two sides.

We are now in a position to consider in how far the above physical forces explain certain physiologic problems.

1. *Is the absorption, into the blood and lymph circulating in the intestinal walls, of substances in solution in the intestinal contents entirely dependent upon the processes of filtration, diffusion and osmosis?* The absorption of weak solutions of highly diffusible substances is probably very largely a matter of osmosis and diffusion, and water passes quickly into the blood because of osmotic attraction, but that other forces ordinarily come into play is very clearly established by the following observations. If a piece of intestine is isolated from the rest by placing two ligatures on it, and the isolated loop filled either with a solution containing the same saline constituents in similar proportions as in blood serum, or better still, with some of the same animal's blood serum, it will be found after some time that all of the solution becomes absorbed into the blood; the contents of the loop are therefore absorbed into the blood, even though the osmotic pressures of the dissolved substances are the same on both sides of the membrane (Weymouth Reid[6]).

The intestinal membrane seems to possess towards readily diffusible

substances a permeability which varies, not at all with the physical diffusibility of the substance, but with its value from a physiologic standpoint. Thus, sodium sulphate and sodium chloride diffuse through ordinary membranes with about equal facility, and yet if a solution containing these two salts is placed in the intestine, the chloride will be absorbed into the blood much more quickly than the sulphate. Sodium sulphate in watery solution diffuses through a membrane fifteen times more quickly than cane sugar, but from the intestinal lumen, cane sugar is absorbed ten times more quickly than sodium sulphate. If. however, the vitality of the epithelium is destroyed, as by first of all bathing it with a solution of sodium fluoride, then the sulphate and chloride will be absorbed at an equal rate.

Although diffusion and osmosis can not therefore play any significant role in the normal process of absorption from the intestine, we must not entirely discount them; under certain circumstances, these physical forces may assert their influence as, for example, when concentrated saline solutions are present. Such solutions will attract water from the blood, and, other things being equal, more will be attracted the less permeable the epithelium happens to be towards the saline employed. Sulphates and phosphates will attract more water than chlorides or acetates. This property of the saline solutions to attract water counteracts the natural tendency for the water to be absorbed, and the large volume of fluid stimulates peristalsis.

2. *Do the physical processes of filtration, diffusion and osmosis suffice to account for the production of urine by the kidneys?* Under normal conditions the molecular concentration of the urine, as determined by the depression of freezing point, is considerably greater than that of the blood. This indicates that excretion must have occurred contrary to the laws of osmosis; in other words, that the renal cells must have compelled dissolved molecules to be transferred from the blood to the urine, although the difference in osmotic pressure would cause them to pass in the opposite direction. This force, sometimes called for want of a better name "vital activity," must depend on the operation of processes that are quite distinct from those of diffusion, etc.; but that they are necessarily of a nonphysical nature (e. g., vital) is less probable than that they depend on some physical process the nature of which our present knowledge does not permit us to understand.

By comparing the osmotic pressures of urine and blood, attempts have been made to measure the work done by the kidney in the production of urine. Thus, it has been found that Δ for normal urine (human) is about 1.8, and for blood about 0.6, from which it may be calculated that in the production of 1 kilogram of urine 150 kilogrammeters of

work are expended.* But that such comparisons of the osmotic pressure of blood and urine are fallacious as an indication of the work of the kidney is evidenced, not alone by the results of the above calculations, but also by the fact that under certain circumstances (as after copious diuresis) the osmotic pressure of the urine may be considerably *lower* than that of the blood. That opposite relationships should exist indicates that differences in osmotic pressure between blood and urine can signify little if anything regarding the work done by the kidney.

For some time after the application of osmotic pressure measurements to the study of biological problems, it was thought that determination of Δ in urine might be of clinical value as a criterion of renal efficiency, especially in one kidney as compared with the other. For this purpose Δ was determined in samples of urine removed from each ureter by catheterization. The tests of renal efficiency based on the rate of excretion of potassium iodide, phenolphthalein, etc., have however been found of much greater value.

3. *Is the formation of lymph purely a physical process?* The osmotic pressure of normal lymph is nearly always somewhat below that of blood serum, although occasionally it has been found to be a trifle higher. Physical processes, such as filtration, might therefore suffice to account for its formation under most conditions. But when we consider the excessive production of lymph that occurs as a result of cellular activity or following the injection of certain substances, called "lymphagogues," it is not so easy to explain the production in such terms, although some interesting attempts have been made to do so by those that are wedded to the mechanistic view. For example, the very marked increase in lymph flow which occurs as a result of muscular exercise or glandular activity has been attributed to the fact that during such processes large molecules become broken down into small ones in the cell protoplasm, so that the osmotic pressure is raised and water is attracted into the the cell until the latter becomes distended and a process of filtration into the neighboring lymph spaces occurs (see page 119).

There are several other physiologic processes of secretion and excretion which might be considered in the present relationship, but the above instances will suffice to illustrate the general principle upon which all of them have to be considered.

*Osmotic pressure corresponding to $\Delta = -0.6°$ C. equals 5,662 mm. Hg (75 m. of H_2O), and that corresponding to $\Delta = -1.8°$ C. equals 16,986 mm. Hg (225 m. H_2O). The difference is therefore equal to a column of water 150 m. high. According to these calculations it would appear that the kidney in producing the average daily output of 1500 c.c. urine performs 225 kilogrammeters of work in comparison with the 14,000 kilogrammeters which the heart is computed to perform in the same time (page 212).

CHAPTER III

ELECTRIC CONDUCTIVITY, DISSOCIATION, AND IONIZATION

The osmotic pressure is not infrequently found to be considerably greater than that expected from the strength of the solution. Although Δ of a gram-molecular watery solution of cane sugar (342 gm. to the liter) is 1.86 (see page 10), that of sodium chloride (58.5 gm. to the liter) is considerably greater. If the hypothesis regarding the relationship of molecular concentration to osmotic pressure is to hold good, it becomes necessary to explain this apparent inconsistency; one must account for a greater number of dissolved units than is represented by the actual number of dissolved molecules (i. e., weight of dissolved substances).

It was observed that the power to conduct the electric current—electric conductivity—in the case of solutions (e. g., of sugar) which have an osmotic pressure that corresponds to the weight of dissolved substances is practically nil, whereas the conductivity of those solutions which give higher osmotic pressure is quite pronounced. Arrhenius made the hypothesis that the conductivity depends on the splitting of molecules into two or more portions or ions, each of which carries either a positive or a negative electric charge, and that it is only when such dissociation occurs that the electric current can be conducted through the solution, the ions serving as it were as floats carrying the electric current. When sodium chloride is dissolved in water, it splits into Na carrying a positive charge and Cl carrying a negative charge, or $Na+Cl-$, as it is written; on the other hand, when sugar is dissolved, the molecules remain unbroken and no electric charges are set free.

Substances which thus dissociate are called *electrolytes,* and those which do not, *nonelectrolytes.* When the electric current is passed through a solution of electrolytes, the ions which carry a positive charge move to the electrode or pole by which the current leaves the solution—that is, in the same directions as the current; and since this electrode is called the cathode, these are called *cations.* Hydrogen and the metals belong to this group. The ions carrying a negative charge go in the opposite direction, against the current—that is, towards the electrode by which the current enters, or the anode; they are therefore called *anions.* They include oxygen, the halogens and the acid groups, such as SO_4, CO_3, etc.

It must be understood that this dissociation into ions is already present in the solution before any electric current passes through it, the ions

being however uniformly distributed throughout—that is, arranged so that the negative charges of the anions precisely neutralize the positive charges of the cations. The electric current causes the electrodes to become charged, the one positively, the other negatively, so that an attractive force is exerted on the ions of opposite sign. This causes the negatively charged ions to migrate towards the positive electrode, and the positively charged, towards the negative electrode. It is this migration of the ions that endows the solution with conducting qualities.

In water, or in a solution of a nonelectrolyte, molecules of H_2O or nonelectrolyte exist thus:

$$\begin{array}{ccccccc}
H_2O & & H_2O & & H_2O & & \\
 & H_2O & & H_2O & & H_2O & \\
H_2O & & H_2O & & H_2O & &
\end{array}$$

In a solution of an electrolyte, the molecules split into ions thus:

$$\begin{array}{ccccccc}
Na^+ & Cl^- & Na^+ & Cl^- & Na^+ & Cl^- & \\
 & Na^+ & Cl^- & Na^+ & Cl^- & Na^+ & Cl^- \\
Na^+ & Cl^- & Na^+ & Cl^- & Na^+ & Cl^- &
\end{array}$$

When an electric current passes through a solution of an electrolyte, the ions tend to arrange themselves thus:

$$\begin{array}{lcccccc}
\text{Cathode}^- & & & & & & \text{Anode}^+ \\
 & Na^+ & Na^+ & Na^+ & Cl^- & Cl^- & Cl^- \\
 & Na^+ & Na^+ & Na^+ & Cl^- & Cl^- & Cl^- \\
 & Na^+ & Na^+ & Na^+ & Cl^- & Cl^- & Cl^-
\end{array}$$

It follows from the above considerations that *the conductivity of a substance in solution will depend on the degree to which it undergoes dissociation.* Furthermore, if we assume that in so far as osmotic pressure phenomena are concerned, each ion behaves in the same way as a molecule, then it follows that the electrical conductivity must be proportional to the extent to which the osmotic pressure is greater than we should expect it to be from the amount of substance actually dissolved.

In the Determination of the Conductivity it is obviously necessary to use standard conditions of depth and width of the fluid through which the current is passed, and to have some standard of comparison. The value is then known as *the specific conductivity,* the standard for comparison being the conductivity of a hypothetical liquid which, if enclosed in a centimeter cube, would offer a resistance of 1 ohm between two opposite sides of the cube acting as electrodes. The actual determination is usu-

ally made in a cylindrical vessel of hard glass (from soft glass enough alkali might be dissolved to affect the results), the electrodes being circular plates of platinum firmly cemented at a known distance from each other (Fig. 5).* This conductivity cell, as it is called, is connected with a suitable electric apparatus for measuring the resistance offered

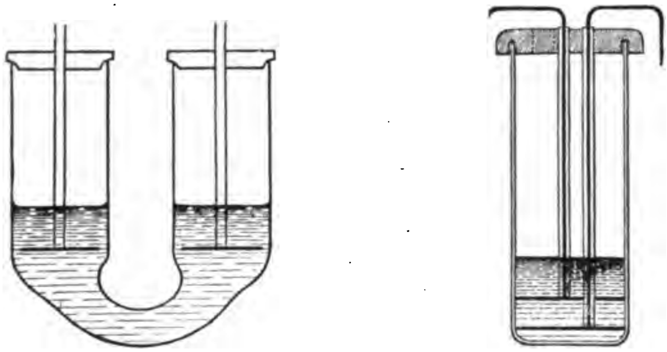


Fig. 5.—Diagram of conductivity cells. The platinum discs are represented by the thick black lines. They are held in position by thick-walled glass tubes, through which they are connected with the terminals by platinum wires. (From Spencer.)

by the solution to the passage of an electric current (Wheatstone Bridge) (see Fig. 6). The resistance is of course inversely proportional to the conductivity.

As a saline solution is progressively diluted, its specific conductivity naturally decreases (since there are now fewer molecules between the

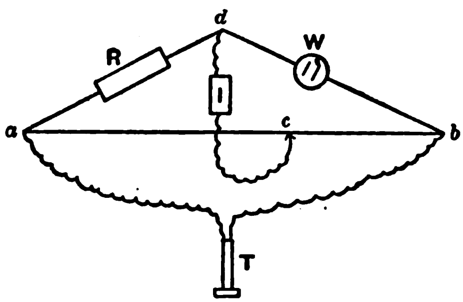


Fig. 6.—Wheatstone Bridge for the measurement of electric resistance: *a-b*, bridge wire; *c*, the movable contact.

two opposite faces of the centimeter cube, and the space between ions or molecules is increased). This result will not, however, tell us whether the salt itself is undergoing any alteration in conducting power as a consequence, for example, of greater dissociation. To ascertain this we must

*This distance is determined not by direct measurement but by calculation from results obtained by testing the actual resistance of a solution whose specific resistance is accurately known.

obtain figures relating to the *same* quantity of salt at each dilution. If we multiply the specific conductivity by the volume of solution in c.c. which contains 1 gram-equivalent (see page 22), a value will be secured which represents the conducting power of a gram-equivalent. This is known as the *equivalent or molecular conductivity,** and is represented by the sign λ. When it is determined for progressively diluted solutions, λ gradually increases, indicating that *the efficiency of the electrolyte itself as a conductor increases with dilution,* because it dissociates more. The extent of this increase is found to become less and less as dilution proceeds. By plotting the values of the molecular conductivity of successive dilutions as a curve, the value at infinite dilution can be ascertained by extrapolation. This value is represented by $\lambda\infty$.

Now, let us see how these facts bear out the theory of electrolytic dissociation. According to this hypothesis the conductivity depends on the number of ions (see page 17), and since it is at a maximum at infinite dilution, *the value $\lambda\infty$ must represent the total number of ions that can be produced by the dissociation of 1 gram-equivalent, and λ that at some other dilution.* If, therefore, we divide λ by $\lambda\infty$ we obtain a value (called a) which must represent the degree to which the electrolyte is ionized at the various dilutions at which λ is measured. From what has been said regarding the osmotic pressure of similar solutions, it is evident that the value a could also be calculated by finding the extent to which the depression of freezing point Δ is greater than would be expected from the number of dissolved molecules. As a matter of fact, it has been found that practically identical values are obtained for many substances, thus furnishing almost incontrovertible proof in support of the dissociation hypothesis. In the cases of weak acids and bases, it is possible to secure a value, called the *dissociation constant* (K), which represents the relative values of a at all dilutions. Since the activity of acids and bases is dependent upon the number of H- and OH-ions, respectively, set free by dissociation, it follows that it must be proportional to K. It will be necessary to postpone a consideration of the application of this constant until we have studied mass action (page 23).

Biological Applications.—The practical value of such knowledge rests, not so much on any direct simple application that can be made of it in explaining physiologic processes, as on the essentially important bearing which it has in enabling us to understand the nature and operation of other physicochemical factors concerned in physiologic processes. Without a clear comprehension of the elemental laws of dissociation, it is impossible to consider such problems as those which concern the activities

*In other words, the molecular conductivity is the specific conductivity divided by the number of gram-equivalents contained in 1 c.c.

of enzymes (mass action, etc.), the occurrence of electric currents during the physiologic activity of muscles, glands, and nerves, and the all-important question of the reaction or H-ion concentration of the body fluids.

Before proceeding to show how these facts concerning the nature of solutions are applicable to the study of physiologic processes, it may be well to indicate one or two instances in which *measurements of electrical conductivity and of dissociation have direct physiologic value.* The circulation time of the bloodflow through an organ can be determined by first finding the electrical resistance of a short piece of the vein of the organ, and then observing the change in resistance which is produced when the conductivity of the blood in the vein is altered by the arrival in it of saline injected into the artery. The interval elapsing between the injection into the artery and the changes in resistance in the vein obviously equals the circulation time (G. N. Stewart).

The same investigator has used measurements by electrical conductivity to study the passage of electrolytes out of the red blood corpuscles into the serum. Under normal conditions the blood serum has a certain electrical conductivity equal to that of a 0.9 per cent sodium-chloride solution. The conductivity of the defibrinated blood is only about one-half that of serum, because it contains corpuscles which are nonconductors and therefore obstruct the free passage of the ions, just as a suspension of quartz powder in a sodium-chloride solution lowers the conductivity of the latter. If anything occurs therefore to occasion a passage of the saline contents of the corpuscles through their walls into the serum, an increase in the electric conductivity will be produced. The value of this method in the investigation of changes in *permeability* of the red corpuscles is dependent on the fact that such migration of electrolytes out of the corpuscles may occur before any of the less diffusible hemoglobin itself has escaped. The rise in conductivity precedes the hemolysis (see page 7).

Although determinations of the specific conductivity of blood and urine under various pathologic conditions have also been made, the results have not been found to possess any diagnostic value or clinical significance. Measurements of the electric conductivity of blood have, however, been used by Wilson[7] and by Priestley and Haldane[8] to detect the degree of dilution when large quantities of water are ingested.

Another application of conductivity measurements in biochemistry has been made in studying the digestive action of proteolytic enzymes (Bayliss). The general action of the enzymes is to break the large undissociated molecules of the higher proteins (albumin, casein, etc.), into smaller molecules (amino acids, etc.), which are partly ionized. As diges-

tion proceeds, therefore, the conductivity of the digestion mixture progressively increases, and is a measure of the rate of digestion.

Applications of the dissociation hypothesis in physiology concern the explanation of such phenomena as the production of electric currents during muscular, glandular, and nervous activity. The exact details of the application are not as yet sufficiently understood to warrant our attempting to do more than indicate the general lines along which the problems are being investigated. Let us, for example, consider how the current of action of muscle may be explained in terms of the dissociation hypothesis. To do so we must delve a little further into physicochemical research, when we shall find that there are two further facts concerning ionized molecules that must be of importance in connection with our problem. The first is that the contribution which each ion makes to the equivalent (or molecular) conductivity of a solution is independent of the other ion with which it is associated; and the second, that ions differ considerably in their conducting power. Since the univalent ions, K., Na., CL′, NO_3', carry charges of the same magnitude,* and yet all do not conduct to the same degree, they must move at different velocities through the solution. We are driven, therefore, to the conclusion that, exposed to the same electric force, different ions have different mobilities; that is to say, when an electric current passes through a solution of an electrolyte, the positively charged ions move towards the cathode at a different rate from that at which the negatively charged ions move towards the anode. Confirmation of this conclusion is obtained by examination of the concentration changes around the two electrodes of an electrolytic cell. The actual velocity of each ion can be determined by experimental means.

*Thus Faraday showed that the amounts of the various ions liberated by electrolysis are in the same ratio as their chemical equivalents.

CHAPTER IV

THE PRINCIPLES INVOLVED IN THE DETERMINATION OF THE HYDROGEN-ION CONCENTRATION

TITRABLE ACIDITY AND ALKALINITY

All acids have one property in common—namely, that they contain hydrogen—and when the acid becomes neutralized, it is this element which becomes replaced by some other cation. Evidently, then, the strength of an acid is proportional to the number of displaceable hydrogen atoms which it contains. It may contain other hydrogen atoms which are so bound up in the molecule that they do not become displaced when an alkali is mixed with the acid. For example, in organic acids like acetic, CH_3COOH, it is only the H atom attached to the COOH group, but not those attached to the CH_3 group, that is replaceable. It must therefore be possible to prepare for every acid a solution having exactly the same neutralizing power as that of any other acid; that is, the same volume of solution must be required in each case to neutralize a given quantity of alkali, the point of neutralization being judged by the change in color of indicators. As a standard a gram-molecular solution of an acid with one displaceable H ion, such as hydrochloric, is chosen. This we call a "normal acid" (N). To prepare a normal solution of acids having two displaceable H atoms, such as H_2SO_4, we can not however use a gram-molecular quantity, but must take one-half of it; and similarly in the case of those with three H atoms, such as H_3PO_4, a one-third gram-molecular solution will be a normal acid solution. For practical purposes, use is very generally made of solutions that are some fraction of the normal, e. g., tenth or decinormal (written N/10), or hundredth or centinormal (N/100).

In a similar way, alkaline solutions can be prepared, a normal alkali being one which exactly corresponds in strength with a normal acid (i. e., can exactly neutralize it). Now, the characteristic of alkalies is that they produce in solution "OH" or hydroxyl ions; so that the process of neutralization must consist in the union of the H ions of the acid with the OH ions of the alkali to form water: $KOH + HCl = KCl + H_2O$. We can, therefore, prepare normal solutions of alkalies by dissolving in 1 liter of water such quantities of alkali (in grams) as will yield the OH required to react with the available hydrogen in normal acid solutions.

Actual Degree of Acidity or Alkalinity.—According to the foregoing method of titration a normal solution of a powerful mineral acid, such as hydrochloric, is no stronger than a normal solution of a weak acid, such as acetic or lactic. It requires no fewer c.c. of N alkali to neutralize it. But the normal solution of the powerful acid tastes more acid, is more toxic, dissolves metals more readily, and in all its other chemical and physiologic properties acts much more quickly than the weak acid, so that the *titrable acidity* or *alkalinity* can not express the real strength of the acid or alkali, or the actual degree of acidity or alkalinity. It is in this connection that the dissociation hypothesis aids us, for it suggests that the degree to which the acid becomes dissociated into H′ and the remainder of the molecule will determine its real strength (see page 16). The question is, how are we to measure the latter? One action of H ions which we may measure is that known as *catalytic*—that is, the power to accelerate reactions, such as the splitting of cane sugar ($C_{12}H_{22}O_{11}$) into glucose and levulose, which otherwise would proceed very slowly (see page 75). If then the real strength of an acid depends on the degree of dissociation which it undergoes, figures representing the catalytic power should correspond with those representing the relative conductivities of the acids in equivalent concentration (see page 19). That this is actually the case is shown in the following table, in which the above values of various acids are given compared with HCl, which is taken as 100.

ACID	CATALYTIC POWER	RELATIVE CONDUCTIVITY
HCl	100	100
Dichloracetic	27	25
Monochloracetic	4.8	4.9
Formic	1.5	1.7
Acetic	0.40	0.42

It will be evident that, if we could measure the concentration of free H ions in a solution—that is, of H ions that are not matched by OH ions—we should have a faithful index of its real acidity. This measurement has been rendered possible by the application of two other physicochemical principles—namely, those of mass action and electromotive force. Since the object of this volume is to present the scientific basis for the various methods that are used in modern medicine, it will be necessary for us to review the main principles of these two actions. We shall see that they apply, not only in the measurement of H-ion concentration, but in many other physiologic processes.

Mass Action

When materials take part in a reaction, some molecules are decomposing while others are being formed. After some time, however, a

condition is reached in which the changes in one direction are exactly offset by those in the other. An equilibrium is said to have become established between the reacting substances. Bearing in mind that the ions and molecules entering into these reactions are constantly moving about and coming in contact with one another, it is easy to see that if we were to add an additional quantity of one kind of molecule or ion, there would be a change all along the line until a new equilibrium was established. If, on the other hand, we were to remove one kind of molecule or ion as fast as it is formed, the equilibrium could never be established, and the reaction would proceed until all of this material had disappeared.

The natural rate at which any chemical reaction proceeds is dependent upon a number of conditions, such as chemical affinity, temperature, catalysis, and concentration. Of these conditions that of concentration is most readily measured. If we maintain all of the conditions other than that of concentration unchanged, and designate this combined influence as K (constant), we shall find that the speed of the reaction is proportional to the molecular concentration of the reacting substances (i. e., the number of gram-molecular weights per liter). In other words, the speed with which two substances, a and b, unite to form other substances, c and d, will be expressed by the equation,

$$k\ (a) \times (b) \rightleftarrows k'\ (c) \times (d);^{*}$$

which means that, when the reaction is complete, the composition of the mixture will be dependent upon the ratio between k and k′. Since however these are both constants, their quotient is also constant (K), and we have the equation, $\frac{(a) \times (b)}{(c) \times (d)} = K$, indicating that no matter how the concentrations a, b, c, and d are varied, reaction will take place in one direction or the other until the concentrations have become adjusted so that K remains unchanged.

As an example of the application of these laws, let us take the reaction which occurs between alcohols and organic acids to form the substances called esters—a reaction which is analogous to that between mineral alkalies and acids to form neutral salts, and which is of special interest to us because it is the reaction involved in the splitting of animal fats. The equation for the reaction is:

$$\underset{\text{(ethyl alcohol)}}{C_2H_5OH} + \underset{\text{(acetic acid)}}{CH_3COOH} \rightleftarrows \underset{\text{(ethyl acetate, an ester)}}{C_2H_5OOCCH_3} + H_2O.$$

Or expressed according to the law of mass action:

$$\frac{[C_2H_5OH] \times [CH_3COOH]}{[C_2H_5OOCCH_3] \times [H_2O]} = K.$$

*The brackets indicate that gram molecular quantities are used.

Now it is clear that if we increase, say, H_2O in the above equation, then in order that K may remain unchanged $C_2H_5OOCCH_3$ must diminish or the substances which form the numerator of the equation must increase, or both these changes must occur. As a matter of fact, in such a case as the above, both of these adjustments take place, for, as the ester breaks down, it must thereby increase the concentration of acid and alcohol. Since in aqueous solutions the reaction occurs in the presence of an excess of water, it is evident that the tendency for an ester in the presence of water is to break down into alcohol and acid, and this must occur in all reactions in the body fluids in which water enters into the equation.

Physiologic Applications.—The application of the law of mass action in the explanation of biochemical processes is very extensive. Most of the reactions which enzymes or ferments are capable of influencing are of the same general nature as that represented above, and the products of their activities are usually the substances on the side of the equation in which no water molecules appear—i. e., they are hydrolytic reactions. Enzymes merely accelerate the reaction (page 72), so that if we start with a mixture of the substances on either side of the equation, all they do is to accelerate the production of a sufficient concentration of those on the other side, until the equilibrium point is reached. For example, an enzyme present in pancreatic juice, called *lipase*, accelerates the breakdown of such esters as neutral fat, which consists of the triatomic alcohol, glycerol, combined with the fatty acids palmitic ($C_{15}H_{31}COOH$), stearic ($C_{17}H_{35}COOH$) and oleic ($C_7H_{33}COOH$):

$$\underset{\text{(the neutral fat, tristearin)}}{C_3H_5\ (O\ OC\ C_{17}H_{35})_3} + 3H_2O \rightleftarrows \underset{\text{(the fatty acid, stearic)}}{3C_{17}H_{35}COOH} + \underset{\text{(glycerol)}}{C_3H_5\ (OH)_3}.$$

Under ordinary conditions the reaction proceeds until nearly all the neutral fat has become decomposed because of the preponderance of water, but if we start with a mixture of fatty acid and glycerol with just enough water to permit the enzyme to act, the reaction will proceed in the opposite direction—i. e., so that some neutral fat will be synthesized. This is called the *reversible action of enzymes.*

Because of the universal presence of water, it is plain that such reversible reactions could not alone be held responsible for the synthesis of neutral fat or of similar substances in the animal body. The only way by which synthesis could occur under these conditions would be if the substance produced along with the water were removed from the site of the reaction as soon as it was formed. This might occur by the precipitation of the substance or by its becoming surrounded by an envelope of some inert material. In the synthesis of neutral fat which

occurs in the epithelium of the intestine out of the fatty acid and glycerol absorbed from the intestinal contents, it is possible that the last mentioned process occurs. In other cases the substance may be carried away by the blood or lymph or urine as fast as it is formed.

The Law of Mass Action as Applied to the Measurement of H-ion Concentration.—Let us now return to the reaction or H-ion concentration of substances in solution. As the standard of neutrality, pure water is chosen. Let us consider, then, how the laws of mass action can be applied in order to enable us to determine the H-ion concentration of pure water. It has been stated above that chemically pure water is incapable of conducting the electric current. This however is not strictly the case, for it conducts to a very slight degree. According to the dissociation hypothesis, it must therefore be represented as containing *molecules* of H_2O and *ions* of $H^\cdot$ and OH', and according to that of mass action there must be a balanced reaction between the molecules and ions represented thus:

$$H_2O \rightleftarrows H' + OH', \text{ or } \frac{[H^\cdot] \times [OH']}{[H_2O]} = K.$$

Since the concentration of $H^\cdot$ and OH' is extremely small, there must always be such an overwhelming preponderance of H_2O molecules that no changes in the concentration of $H^\cdot$ and OH' will be capable of appreciably affecting the concentration of H_2O; in other words, one may omit the denominator of the equation and write it $[H^\cdot] \times [OH'] = K$. If then we know the value of K, it will only be necessary to measure the concentration of *either* $H^\cdot$ or OH' in order to express in numerical terms the reaction of the solution. It has been found that the value of K is about 1×10^{-14},* and since the concentrations of $H^\cdot$ and OH' are necessarily equal in pure water, it follows that $[H] = [OH] = \sqrt{1 \times 10^{-14}}$, i. e., each ion has a concentration of 1×10^{-7}. This means that water contains approximately 1 gram mol. each of $H^\cdot$ and OH' ions, or 1 gram $H^\cdot$ and 17 grams OH' ions, in 10^{-7} or 10,000,000 liters. A consequence of the above law is that no matter how much the concentration of one ion is increased by adding another substance, the solution must still contain some of the other ion. Thus, in acid solutions *con.* $H^\cdot$ must increase and *con.* OH' must decrease in such proportion that the two multiplied together equals about 1×10^{-14}. *The H-ion concentration can be used therefore to express the reaction of neutral, acid and alkaline solutions.*

In place of water, let us substitute decinormal hydrochloric acid

*The sign 10^{-14} is simply a convenient way of expressing the degree of dilution. It gives the number of times the value standing in front of it must be multiplied by 10 in order to find the degree of dilution.

(0.1 *N* HCl)—that is, a hydrochloric acid solution containing one tenth of the molecular weight of hydrochloric acid in grams dissolved in a liter of water. At this dilution HCl is 91 per cent dissociated; therefore the H-ion concentration (or C_H as it is written for short) is 0.091 *N*, or, in mathematical notation, 9.1×10^{-2}.

Method of Expressing C_H.—To avoid the necessity of having to use several figures to express C_H, as has been done above, Sörenson has introduced a scheme by which only one figure is required. This figure, designated by P_H, is found by subtracting from the power of ten (i. e., the figure standing behind 10) the common logarithm of the figure expressing the normality of the acid.* In a decinormal HCl solution, therefore, we must subtract from the power 2, the common log. of 9.1, which is .96 (ascertained from logarithm tables), leaving 1.04. . Take another example: decinormal acetic acid is dissociated only to the extent of 1.3 per cent; C_H is therefore 0.0013 normal, or 1.3×10^{-3}. Since the logarithm of 1.3 is .11, P_H equals $3 - .11$, or $^-2.89$.†

The only objection to the use of the exponent P_H as an expression of the H-ion concentration is that it increases in magnitude as the acidity becomes less; this is because the negative sign of the power is disregarded. As stated above, it is usual to express the strength of alkalies as well as acids in terms of C_H, or P_H, because it is easier to measure the concentration of H ions than of OH ions. A 0.1 NaOH solution is 84 per cent dissociated; therefore the "OH" ion is 0.084 *N* (i. e., 0.084 gram equivalents OH per liter), and since the product of the H· and OH′ concentrations must always equal $10^{-14.14}$ (at 20° C.), it is clear that as the H ion increases in concentration, the OH ion must reciprocally decrease. Expressed according to the above scheme, the 0.084 *N* NaOH solution gives P_H 13.06; thus, $0.084 = 8.4 \times 10^{-2}$; the log. of 8.4 is .924, and this subtracted from the power $-2 = 1.08$ as P_{OH}, or $14.14 - 1.08 = 13.06$ as P_H.**

Similarly, P_H of 0.1 *N* NH_4HO solution is 11.286. Its dissociation is 1.4 per cent; therefore the solution contains only 0.0014 gram equivalents HO—i. e., 1.4×10^{-3} $P_{OH} = 3 - 0.146 = 2.854$ $\therefore$ P_H $14.14 - 2.854 = 11.286$.‡

*Strictly speaking, P_H is the logarithm to the base 10 of the concentration of H ions in grams per liter, the negative sign being understood.

†If we wish to express the value of P_H in ordinary notation, we must find the antilogarithm of the difference between the value of P_H and the next higher whole number; e. g., if $P_H = 7.45$, the antilogarithm of 0.55 being 3.55, the C_H is 3.55×10^{-8}, or 0.000,000,0355 *N*, or 3.55 gm. mol. H ion in 100,000,000 liters.

**It must be remembered that the power of a number indicates the number of times by which that number must be multiplied by ten; thus, $P_H{}^{-6}$ does not mean that the H ion is six times less than $P_H{}^{6}$, but 1 x 10 x 10 x 10 x 10 x 10 x 10, or 1,000,000 times less. Similarly, $P_H{}^{-3}$ is 1000 times as great as $P_H{}^{-6}$, not twice as great.

A solution containing almost exactly 1 gram molecule of dissociated hydrogen in 10,000,000 liters constitutes a neutral solution ($P_H = 7$).

‡The expressions P_H and C_H may be used indiscriminately, but when the numerical value is given, it is most convenient to use the former.

Application of the Law of Mass Action in Determining the Real Strength of Acids or Alkalies.—We have seen that it is the degree of dissociation upon which the real strength of an acid depends and that this varies with dilution (page 19). The equilibrium between the undissociated and dissociated molecules may therefore be shifted in either direction by changing the concentration; in other words, the process of dissociation is a reversible reaction, and may be represented as $AB \leftrightarrows A' + B^{\cdot}$. The law of mass action must apply in such a case, and as a matter of fact it has been found that a constant can be calculated, which is known as the dissociation constant.* It is an expression of the inherent ability of the acid to dissociate into ions, and is therefore the best measure of the strength of the acid. This is strictly the case for all of the weaker acids, but strong mineral acids (and bases) do not give a satisfactory constant, so that the comparison must not be made between them and weaker ones. That the dissociation constant expresses the relative strength of organic acids can be shown by comparing its value with that of the rate at which cane sugar is inverted (see page 23), this being proportional to the concentration of the H ions present. K for some organic acids is: Acetic, 0.000018; Formic, 0.000214; Benzoic, 0.00006; Salicylic, 0.00102.

*The equation is $\frac{a^2}{(1-a)V} = K$, where it is supposed that in volume V of the solution there is 1 gram-equivalent of electrolyte, and that the degree of dissociation is a; the quantity of undissociated electrolyte stated in a fraction of a gram-equivalent will be 1-a, and the quantity of each ion a. To illustrate, let us take acetic acid in various dilutions:

V	a	$K \times 10^5$
0.994	0.004	1.62
2.02	0.00614	1.88
15.9	0.0166	1.76
18.1	0.0178	0.78

CHAPTER V

THE PRINCIPLES INVOLVED IN THE MEASUREMENT OF THE HYDROGEN-ION CONCENTRATION (Cont'd)

THE METHODS OF MEASUREMENT

The Electric Method

In order to understand the principle of the standard method used for measuring the H-ion concentration, it is necessary that a few words be said concerning the factors governing **the development of electric currents** in chemical batteries. There may be a further application of this knowledge in connection with the generation of the electric currents which occurs during physiologic activity, as in active glands and muscles.

When a metal is immersed in a solution of one of its salts, it has a tendency to give off ions into the solution. Similar ions are, however, already present in this solution, and these, by their osmotic pressure, tend to oppose the passage of the ions coming from the metal. The force with which the metal sends out its ions into the solution is called *the electrolytic solution pressure.* If this is equal to the osmotic pressure of the metallic ions in the solution, there will be no electric current generated, but if it is greater or less than the osmotic pressure of the metallic ion, an electric current will be set up. When the solution pressure is the greater, the metal will become negatively charged, because its ions carry positive charges into the solution (cations); on the contrary, when the osmotic pressure is greater than the solution pressure, the metal will have a positive charge, owing to the receipt of the positive cations from the solution.

Because of a force called *electrostatic attraction,* the ions given off from the metal can not travel any measurable distance from the oppositely charged mass of metal, so that from one of the electrodes alone it is impossible for us to lead off any electric current. For this purpose we must form a circuit. This is done in the manner shown in Fig. 7 by connecting side tubes coming from the electrode vessels with an intermediate vessel containing a solution of high conductivity and by connecting the metals by wires. If the circuit is formed between the same metals in solutions of the same concentration, no electric current will be generated, because the two electrode potentials will be

equal and in opposite directions to each other. On the other hand, should the concentration of the metallic ion in the solutions be unequal, the electromotive force will flow from the one electrode to the other, and the pressure with which it flows will be equal to the difference in concentration of the two solutions. This is the principle of a *concentration cell,* and if we know the concentration of one of the solutions composing it, and then proceed to measure the electromotive force, we can obtain the concentrations of the other solution by difference. To do this we must employ a formula which takes into consideration the relation between the potential and the concentration of the solution.

The potential of an unknown electrode composed of a metal in contact with a solution of one of its salts may also be determined by making it one pole of a battery of which the other pole is composed of a standard electrode of unchanging known potential. An electrode of the latter

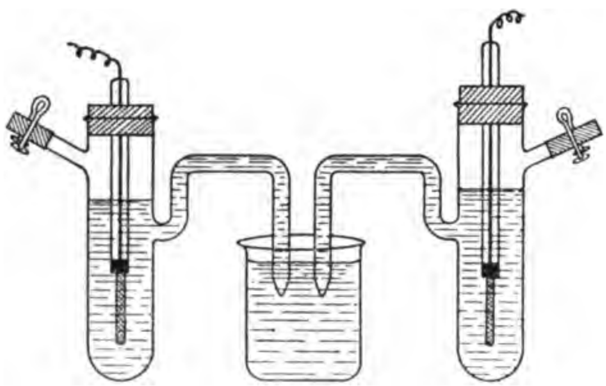


Fig. 7.—Diagram to show type of electrodes used in studying electromotive force. The metal in each electrode is connected (through a glass tube) with a platinum wire, to which the apparatus for measurement of the voltage is connected. The metal dips into a solution contained in the electrode vessel and filling the side tube. The latter dips into an intermediate vessel containing saturated KCl solution. The currents flow through the circuit under the following conditions: (1) dissimilar metals dipping into the same fluid; (2) similar metals dipping into different fluids; (3) dissimilar metals dipping into different fluids.

type can most readily be made by bringing pure mercury in contact with a saturated solution of calomel (Hg_2Cl_2) in normal potassium chloride solution. Under suitable conditions (i. e., when the circuit is completed), a potential of +0.560 v. is developed in this so-called *calomel electrode**—that is, positive ions of mercury are deposited on the mercury from the calomel solution at this pressure. Suppose that we connect a calomel electrode, through the intermediation of some solution which

*The calomel electrode consists of a suitably shaped glass vessel containing pure mercury, connected by means of a platinum wire with a conductor, and filled with a saturated solution of pure mercurous chloride in normal KCl solution up to such a level that it also fills a side tube connected with a vessel containing a saturated solution of potassium chloride. Into this vessel also runs a similar side tube from the unknown electrode. By having an excess of undissolved calomel in the solution in the calomel electrode its saturated condition is maintained during the chemical changes which accompany the production of the electric current.

will serve as a good conductor, with another electrode, the two electrodes being also connected by wires with electrical apparatus for measuring the total potential of the battery; then by adding $+0.560$ v. to or subtracting this value from the total potential (depending on the sign of the unknown electrode) we can tell the potential of the unknown electrode.

We have discussed these principles of electrochemistry because they form the basis upon which depends the standard method for **the determination of the H-ion concentration of fluids.** Suppose, for example, that in place of using a metal in the construction of one electrode, we use an electrode consisting of a layer of pure hydrogen gas in contact with a solution in which are free H ions; then the rate at which H ions

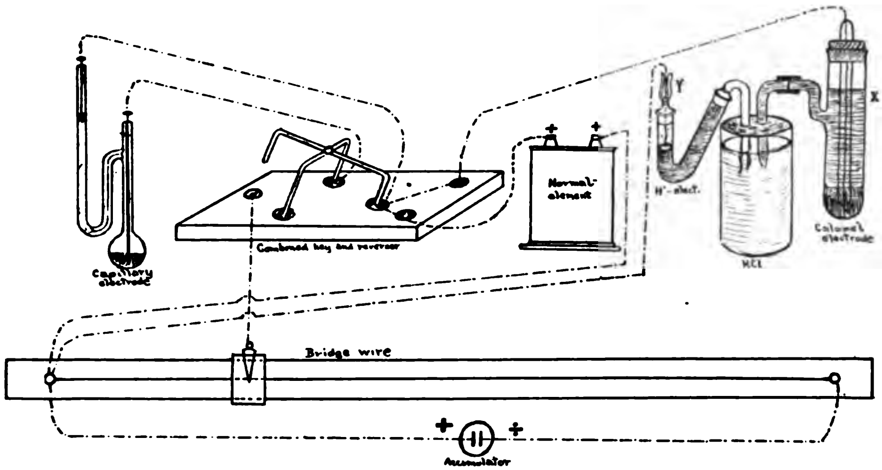


Fig. 8.—Diagram of apparatus for the measurement of the H-ion concentration. The current generated in the battery (composed of calomel electrode, connecting vessel with KCl solution, and the H electrode) or that from the normal element is transmitted through a reversing key to the bridge wire, where the voltage is compared with a steady current flowing through the bridge wire from an accumulator. The capillary electrometer is used to detect the flow of current at various positions of the movable contact on the bridge wire. (Modified from Sörensen.)

become added to the solution from the H layer, or taken from it, will depend on the concentration of H ions in solution. In order to secure a hydrogen electrode fulfilling the above requirements, it is necessary to employ some means by which a layer of hydrogen may be furnished, and fortunately this can be done by taking advantage of the property which spongy platinum possesses of absorbing large quantities of this gas. It is also necessary to keep an atmosphere of pure H in contact with the fluid.

As is the case of the simpler cells described above, there are two types which we might use for measuring the electromotive force generated in the unknown electrode: a concentration cell composed of two

hydrogen electrodes, of which one contains a solution of *known* H-ion concentration, and the other the solution in which this is unknown; and a cell of which one electrode is a standard calomel electrode and the other, a hydrogen electrode containing the unknown solution.

The exact arrangement of the apparatus in which the calomel electrode is used will be seen in the accompanying sketch. The hydrogen electrode, it will be noticed, is a very small V-shaped tube, in which is suspended a platinum wire coated with spongy platinum and dipping into a solution which nearly fills the tube. The space above the solution is filled with pure hydrogen. This and the calomel electrode are connected with suitable electric measuring instruments, and the circuit is completed by connecting the two electrodes by means of an intermediate vessel containing a saturated solution of potassium chloride. This connecting solution is used because it has been found that the electric currents set up at the contact between different solutions are so small that they can be disregarded.*

As outlined above, the hydrogen electrode is that which is used to determine the H-ion concentration of blood, the particular point about it, in comparison with the apparatus used for simpler solutions, being that the hydrogen is not changed in the course of the experiment. This precaution is to prevent the carbon dioxide of the blood from being "washed out" of it by a frequently changing atmosphere of hydrogen. Many inaccuracies in the earlier results obtained by this method were due to the removal of carbon dioxide, which, as we shall see later, is one of the chief acids contributing to the H-ion concentration of blood.

The Indicator Method

As pointed out in a previous chapter (page 22), the method of titration for acidity or alkalinity in which a standard solution of alkali or acid is added until a certain change in the color of a suitable indicator is detected, does not afford any information regarding the H-ion concentration actually present in the solution. It tells us the *total concentration* of available acid or base, both dissociated and undissociated. By modification of the method of procedure, however, we may also use indicators for determining the H-ion concentration. The principle of this method depends on the fact that there are certain dyes which change quite distinctly in tint with very slight changes in the H-ion concentration, so that if we use dyes which possess this property at a point near that of neutrality (i. e., between $P_H6.5$ and P_H8), we can es-

*A description of the technic for measuring the electric potential developed by the cell would be out of place here. Suffice to say that the strength of the current is compared with that of a current of known strength furnished by a normal cell, the comparison being made by a bridge wire F, a capillary electrometer H being employed to detect the direction and degree of current.

timate the H-ion concentration of the body fluids with very remarkable accuracy, provided certain precautions are taken to circumvent the disturbing influence which the protein and salts in these fluids may have on the color change.

To understand this use of indicators, it is important to bear in mind that one solution reacting neutral to one indicator may have a H-ion concentration which differs very markedly from that of another solution reacting neutral to another indicator. This is because indicators react to different H-ion concentrations. A solution that is neutral to phenolphthalein has a P_H of about 9, whereas one neutral to methyl orange has a P_H of about 4. This can be very clearly shown by titrating a solution of phosphoric acid with decinormal alkali. After a certain amount of alkali has been added it will be noticed that methyl orange changes from red to yellow, but after it has changed and is therefore alkaline as judged by this indicator, it still remains distinctly acid towards phenolphthalein (shows no red tint) even though considerably more alkali is added. The methyl orange is, therefore, itself unresponsive to weak acids such as remain after the greater part of the phosphoric acid has been neutralized by the alkali.

The series of indicators which has been employed for this purpose is given in the accompanying table, along with the P_H limits through which they change in color.

List of Indicators

Chemical name	Common name	Concentration	Color change	Range P_H
		per cent		
Thymol sulfon phthalein (acid range)	Thymol blue	0.04	Red-yellow	1.2-2.8
Tetra bromo phenol sulfon phthalein	Brom phenol blue	0.04	Yellow-blue	3.0-4.6
Ortho carboxy benzene azo di methyl aniline	Methyl red	0.02	Red-yellow	4.4-6.0
Ortho carboxy benzene azo di propyl aniline	Propyl red	0.02	Red-yellow	4.8-6.4
Di bromo ortho cresol sulfon phthalein	Brom cresol purple	0.04	Yellow-purple	5.2-6.8
Di bromo thymol sulfon phthalein	Brom thymol blue	0.04	Yellow-blue	6.0-7.6
Phenol sulfon phthalein	Phenol red	0.02	Yellow-red	6.8-8.4
Ortho cresol sulfon phthalein	Cresol red	0.02	Yellow-red	7.2-8.8
Thymol sulfon phthalein (see above)	Thymol blue	0.04	Yellow-red	8.0-9.6
Ortho cresol phthalein	Cresol phthalein	0.02	Colorless-red	8.2-9.8

(W. M. Clark and H. A. Lubs.)[9]

These dyes may now be obtained in this country.

Briefly stated the method for measuring the H-ion concentration consists in preparing a series of solutions containing known concentrations of H-ion—that is to say, of known P_H—and adding to each solution an equal amount of an indicator which exhibits easily distinguishable changes in tint at H-ion concentrations approximating those believed to be present in the unknown solution. The same indicator is added to the unknown solution, which is then placed side by side with the standards to find with which of them it most closely matches. The series of solutions of known H-ion concentration is prepared by mixing fifteenth normal solutions of Na_2HPO_4 and KH_2PO_4 in varying proportions as given in the following table:

PREPARATION OF STANDARD SOLUTIONS

The solutions are mixed in the proportions indicated below to obtain the desired P_H:*

P_H	6.4	6.6	6.8	7.0	7.1	7.2	7.3	7.4	7.5	7.6	7.7	7.8	8.0	8.2	8.4
Primary Potas. Phos., c.c.	73	63	51	37	32	27	23	19	15.8	13.2	11	8.8	5.6	3.2	2.0
Secondary Sodium Phos., c.c.	27	37	49	63	68	73	77	81	84.2	86.8	89	91.2	94.4	96.8	98.0

(From Levy, Rowntree and Marriott.)

*Standard phosphate mixtures are prepared according to Sörensen's directions as follows:

1/15 mol. acid or primary potassium phosphate.—9.078 grams of the pure recrystallized salt (KH_2PO_4) are dissolved in freshly distilled water and made up to 1 liter.

1/15 mol. alkaline or secondary sodium phosphate.—The pure recrystallized salt ($Na_2HPO_4.12H_2O$) is exposed to the air for from ten days to two weeks, protected from dust. Ten molecules of water of crystallization are given off and a salt of the formula $Na_2HPO_4.2H_2O$ is obtained; 11.876 grams of this are dissolved in freshly distilled water and made up to 1 liter. The solution should give a deep rose red color with phenolphthalein. If only a faint pink color is obtained, the salt is not sufficiently pure.

The indicator method is extremely accurate when used with pure solutions of acids, but, as mentioned above, it is apt to be inaccurate, at least with most indicators, when protein or inorganic salts are present in the solution, and of course it is quite unusable with colored fluids such as blood. In order to overcome these difficulties, the *dialysis method* has recently been evolved. It consists in placing the fluid—blood, for example—in a dialyser sac composed of celloidin and about as large as a small test tube. The sac is placed in a wider test tube of hard glass containing an isotonic solution of sodium chloride that has been carefully tested to ascertain that it is strictly neutral. The amount of blood or serum required for this method is only 2 or 3 c.c., and the amount of salt solution placed outside the sac should be about the same. It takes only from five to ten minutes for dialysis to occur. The celloidin sac is then removed, a few drops of the indicator are thoroughly mixed with the dialysate, and the tube compared with the series of standards until the corresponding tint is matched. This indicates the H-ion concentration in the dialysate. The tints produced by using sulphonephenolphthalein are reproduced as nearly as possible

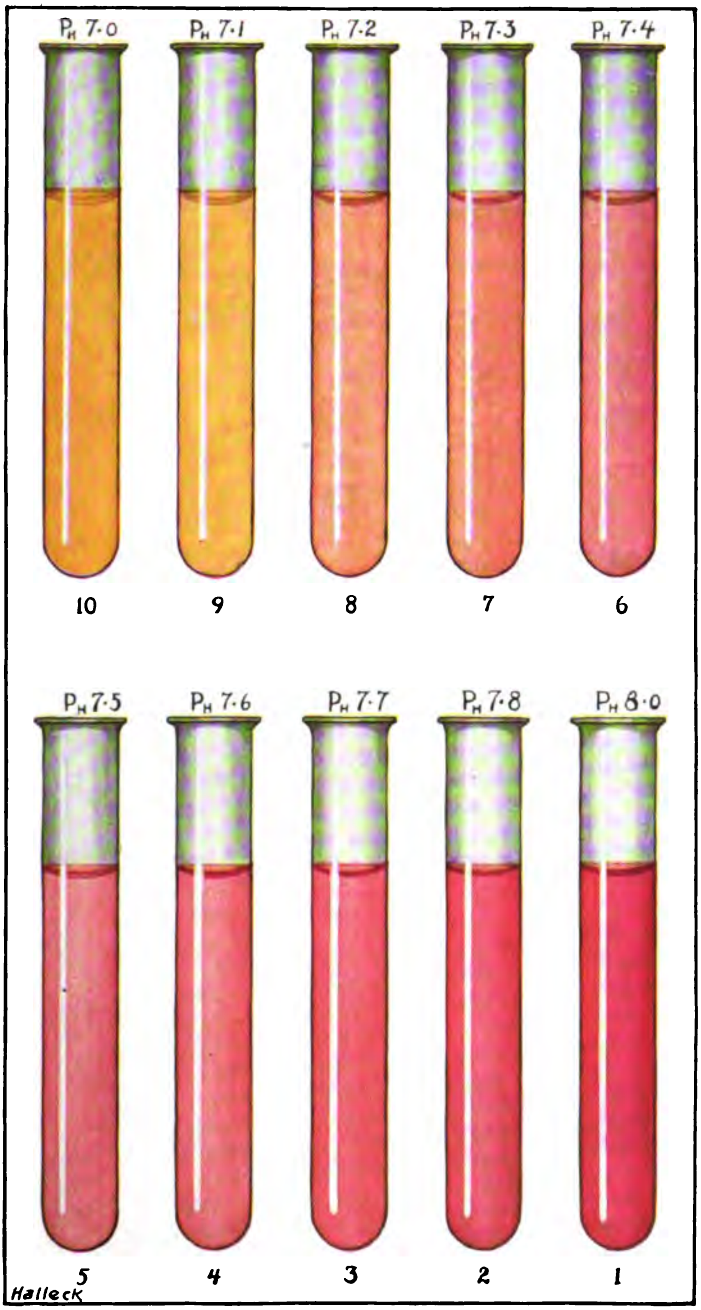


Fig. 9.—Chart showing approximately the tints produced by adding sulphophenolphthalein to a series of phosphate solutions of the H-ion concentrations indicated in each case by P_H.

in the accompanying chart. The H-ion concentration of the unknown solution is that of the tint with which it matches in the series.

It might be thought that this method would be inaccurate because of the loss of carbon dioxide from the blood. By actual experiment, however, it has been found that, if the blood is collected with certain precautions, the error is negligible. The method is, therefore, a most useful one clinically.

The following table gives the hydrogen-ion concentration or true reaction of the body fluids.

FLUID	PH	FLUID	PH
Blood	7.4	Muscle juice (fresh)	6.8
Urine	6.0	Muscle juice (autolyzed)	Variable
Saliva	6.9	Pancreas extract	5.6
Gastric juice (adult)	0.9-1.6	Peritoneal fluid	7.4
Gastric juice (infant)	5.0	Pericardial fluid	7.4
Pancreatic juice (dog)	8.3	Aqueous humor	7.1
Small intestinal contents	8.3	Vitreous humor	7.0
Small intestinal contents (infant)	3.1	Cerebrospinal fluid	7.2
Bile from liver	7.8	Cerebrospinal fluid	8.3
Bile from gall bladder	5.3-7.4	Amniotic fluid	7.1
Perspiration	7.1	Amniotic fluid	8.1
Perspiration	4.5	Milk (human)	7.0-7.2
Tears	7.2	Milk (cow)	6.6-6.8
		Milk (goat)	6.6
		Milk (ass)	7.6

(W. M. Clark and H. A. Lubs.)

CHAPTER VI

THE REGULATION OF NEUTRALITY IN THE ANIMAL BODY AND ACIDOSIS

Nothing is more constant in the animal economy than the H-ion concentration (C_H) of the fluids which bathe the tissues. This regulation is fundamentally of a physicochemical nature, depending on the interaction of alkalies with acids, of which carbonic and phosphoric acids and the proteins are the most important.* When different amounts of acids or alkalies are added to water, the range of variation in H ion is very extensive, whereas in blood the range is very limited indeed, not extending beyond P_H7 and $P_H7.52$ (i. e., C_H never goes above that of a 0.000,000,1 *N* solution or below that of a 0.000,000,03 *N* solution). In other words blood can withstand considerable additions of acid or alkali without much change.

Buffer Substances.—The chemical reactions upon which this remarkable constancy in reaction depends have been explained by Lawrence J. Henderson.[10] The fundamental equations are as follows:

$$M_2HPO_4 + HA = MH_2PO_4 + MA, \text{ and}$$
$$MHCO_3 + HA = H_2CO_3 + MA,$$
when M = a basic radicle, and A, an acid radicle.

Now it has been discovered that weak acids, like carbonic and phosphoric, possess the remarkable property of maintaining the reaction constant when they are present in a solution which also contains an excess of their salts. Under these circumstances the concentration of ionized hydrogen is almost exactly equal to the product of the dissociation constant† of the acid (see page 19) multiplied by the ratio between free acid and salt; in other words,

$$H^{\cdot} = K \times \frac{[HA]}{[BA]}.$$

If carbonic acid is present in a solution of bicarbonates so that there

*According to circumstances, proteins may act either as acids or as alkalies. They are therefore called amphoteric.

†The ionization constant has already been referred to as a figure which expresses the tendency of a weak acid or base to dissociate in an aqueous solution. "It expresses the proportion in which the nondissociated part is capable of existing in the presence of its ions," and therefore is a gauge of the strength. The dissociation constant amounts to about 0.000,000,5 for carbonic acid; that is, the dissociation of H_2CO_3 into $H^{\cdot} + HCO_3'$ at room temperature will be such that the concentration of H-ion equals a 0.000,000,5 N solution.

are equivalent quantities of *free* H_2CO_3 and bicarbonate—i. e., $\frac{[HA]}{[BA]}=\frac{1}{1}$ —the H-ion concentration will be exactly the same as the dissociation constant of carbonic acid; therefore 0.000,000,5 N ($P_H = 6.31$), or about five times the value of neutrality, 0.000,000,1 N ($P_H = 7.31$). If ten times as much free carbonic acid as bicarbonate is present, then the H-ion concentration will be fifty times that of neutrality, i. e., $\frac{[HA]}{[BA]}=\frac{10}{1}$ $\times$ 0.000,000,5 = 0.000,005 ($P_H = 5.31$); if there is ten times less carbonic acid than bicarbonate, the H-ion concentration will be one-half that of neutrality, i. e., $\frac{[HA]}{[BA]}=\frac{1}{10}\times 0.000,000,5 = 0.000,000,05$) ($P_H = 7.31$); or if twenty times less, one fourth ($P_H = 7.6$). Since a large amount of bicarbonate is actually present in blood (enough to yield from 50 to 65 c.c. CO_2 per 100 c.c. of blood) (see page 391), and the free carbonic acid undergoes fluctuations which are only trivial when compared with those which have been chosen in the above examples, it is clear that there must be very little change in the H-ion concentration of the blood in comparison with the variations which would occur were no bicarbonate present.

Another weak acid which acts like carbonic in maintaining neutrality is acid phosphate (MH_2PO_4), and for the same reason—namely, that its dissociation constant is of similar magnitude to the H-ion concentration. Although the blood plasma itself contains much less phosphate than bicarbonate, the tissues contain a considerable amount, which enables them to maintain their neutrality. This action of bicarbonates and phosphates is styled the buffer action, meaning that it serves to damp down the effect on the H-ion concentration which additions of acids or alkalies would otherwise have. As pointed out by Bayliss, however, a better word to use would be "tampon action," since the substances actually soak up much of the added $H^{\cdot}$ or OH' ions. It is not confined to the fluids of the higher animals, but is very widely distributed throughout nature; for example, in the ocean and in the fluids of marine organisms and animalcules (see L. J. Henderson).[11]

Although the actual reaction by which neutrality is maintained is purely of a physicochemical nature, some provision must obviously be made so that the acid and basic substances that take part in it may be supplied and those produced by the reactions removed as occasion requires. The *source of supply* is partly exogenous and partly endogenous. The exogenous source is the basic and acid substances present in the food; and although we do not ordinarily attempt to control the amounts of these substances ingested, we may do so, as, for example, by the persistent administration of soda in cases of pathologic acidosis. The endogenous source depends on the constant production in metabolism

of acids such as carbonic, phosphoric, lactic, and sulphuric, and of alkalies such as ammonia. Amphoteric substances, like amino acids and proteins, may functionate either as acids or as alkalies. Whatever may be its source, a considerable reserve of alkali is undoubtedly available in the animal organism. This required store of alkali appears to be automatically liberated as a result of the physicochemical process.

The removal is affected by three pathways: (1) through the lungs gaseous carbonic acid is eliminated; (2) through the kidneys, the fixed acids; and (3) through the intestines, some of the phosphoric acid.

Carbonic acid is produced in large amounts in the normal process of metabolism, and is excreted in a gaseous condition by the lungs. Variation in its excretion is the most important mechanism for controlling *temporary changes* in C_H. In order to make this clear, it may be well to revert for a moment to the physicochemical equation by which carbonic acid is enabled to maintain neutrality. This may be written: $C_H =$ molecular ratio $\frac{H_2CO_3}{NaHCO_3}$. The ratio may be increased either by adding free carbonic acid to the blood (as by causing an animal to respire some of the gas), or by the addition of some other acid (e. g., oxybutyric, as in diabetes) which will decompose some of the $NaHCO_3$ and produce H_2CO_3. The increase which these changes would cause in C_H of the blood is prevented by the remarkable sensitivity of the respiratory center to changes in C_H. An increase which is much less than can be measured by physicochemical means stimulates the center, causing increased pulmonary ventilation, so that the carbonic acid is immediately eliminated through the lungs. This elimination does not stop when the old level of carbonic-acid concentration is reached, but proceeds until the original ratio $\frac{H_2CO_3}{NaHCO_3}$ is again attained in the blood, and C_H is restored exactly to its original value. If it stopped at the old CO_2 concentration, the ratio would be too high because there is less $NaHCO_3$.

THE THEORY OF ACIDOSIS

Although these considerations indicate that variations may occur in the bicarbonate content of the blood without any significant change in C_H, they also show that the bicarbonate content must be a criterion of the acid-base balance of the blood, and probably of the body fluids in general. As pointed out by Van Slyke,[12] bicarbonate represents the excess of base which is left over after all the *fixed* acids have been neutralized. It represents the base that is available for the neutralization of any excess of such acids that may appear—a measure of the reserve of "buffer substance" or, more specifically, *the alkaline reserve* of the body.

Under normal conditions the amount of $NaHCO_3$ in blood plasma is very constant (amounting to 50-65 vols. per cent CO_2), and when it is reduced, it indicates that an excess of fixed acid must be present. This is taken by Van Slyke and others to constitute the real definition of *acidosis*—namely, "a condition in which the concentration of bicarbonate in the blood is reduced below the normal level." If the respiratory center for any reason should not respond promptly enough to an increase in the molecular ratio $\frac{H_2CO_3}{NaHCO_3}$, and C_H consequently become greater, the condition is called *uncompensated acidosis,* but if the center does respond so that C_H is held constant (although $NaHCO_3$ is decreased), the condition is one of *compensated acidosis.*

For practical reasons, therefore, the study of pathologic acidosis depends on an estimation of the bicarbonate content of the blood or, since it is simpler to carry out and is of equal value, of the plasma. When plasma is obtained by removing blood from a vein of the arm and centrifuging immediately out of contact with air (so that CO_2 may not be lost from it) it contains approximately 60 vols. per cent of CO_2. Since we know that the partial pressure of CO_2 in blood is equal to 42 mm. Hg (ascertained from determinations of the alveolar CO_2) (see page 344), we can calculate how much of the 60 vols. per cent must be in simple solution by application of the law of solution of gas in a liquid (page 336). It has been found that plasma at body temperature and at 760 mm. Hg (atmospheric pressure) dissolves 0.54 per cent CO_2, so that at 42 mm. it will dissolve $\frac{42}{760} \times 100 \times 0.54 = 3$ vols. per cent. Transcribing the figures to our equation we get $\frac{[H_2CO_3]}{[NaHCO_3]} = \frac{3}{60}$, or $\frac{1}{20}$.*

This definition of acidosis leaves out of regard all conditions that may raise the ratio $\frac{H_2CO_3}{NaHCO_3}$ by the addition of H_2CO_3 without decomposing any of the $NaHCO_3$, such, for example, as occurs when an excess of free carbonic acid is present in the blood plasma. Since increases in free CO_2 are not infrequent in both health and disease—e. g., asphyxial conditions—the above definition is not sufficiently comprehensive. When we come to study the control of the respiratory center, we shall see that an increase in the ratio $\frac{H_2CO_3}{NaHCO_3}$ of sufficient magnitude to cause an actual increase in C_H can be produced by causing an animal to respire air

*This agrees sufficiently with the result as calculated from the known values of the equation $\frac{H_2CO_3}{NaHCO_3} = \frac{\lambda C_H}{K}$. Thus, if we take C_H as 0.35×10^{-7}, λ as 0.605 for blood conditions, and K as 4.4×10^{-7} (Michaelis and Rona), we get $\frac{H_2CO_3}{NaHCO_3} = \frac{0.605 \times 0.35 \times 10^{-7}}{4.4 \times 10^{-7}} = \frac{1}{21}$

containing an excess of CO_2—a true acidosis, but one for which no place is found in the above definition.

Nevertheless, Van·Slyke's definition has a real value, because it emphasizes the importance of a determination of the bicarbonate as a criterion of the degree of the forms of acidosis usually met with in disease. The bicarbonate under such conditions may become reduced either because of the appearance of improperly oxidized fatty acids, like β-oxybutyric and acetoacetic, when carbohydrate metabolism is upset as in diabetes or starvation, or because the acids produced by a normal metabolism are inadequately eliminated by the kidneys, as in nephritis.

Accordingly, if the respiratory mechanism and increased mass movement of the blood (for an increase in C_H accelerates this also) should fail to eliminate CO_2 quickly enough so as to keep the $\frac{H_2CO_3}{NaHCO_3}$ ratio at one twentieth, then C_H will rise. This is not likely to happen until a large part of the $NaHCO_3$ has been used up, so that an estimation of that actually present must be a reliable index of the proximity to this condition.

A sustained increase in C_H is incompatible with life. The $NaHCO_3$ is the buffer, the factor of safety which prevents its occurrence. Although it is only in arterial blood (i. e., after elimination of excess of CO_2 by the lungs has been accomplished) that constancy in the ratio $\frac{H_2CO_3}{NaHCO_3}$ can be expected, it is fortunate, for practical reasons, that venous blood collected during muscular rest and without stasis should be only slightly different.

When acids are added to the blood, they will first of all be neutralized by the "buffers" of the plasma—namely, $NaHCO_3$ and protein, as we have seen. But this is only the first line of defense against acidosis, for buffer substances present in the corpuscles may also be used. This intracorpuscular reserve of alkali is mobilized partly by transference of K and Na from corpuscle to plasma, but mainly by that of HCl from the plasma into the corpuscle, so releasing base in the former to combine with the added acid (e. g., H_2CO_3), according to the equation: $H_2CO_3 + NaCl \rightleftarrows NaHCO_3 + HCl$. The HCl on entering the corpuscle reacts with phosphates according to the equation: $HCl + Na_2HPO_4 \rightleftarrows NaH_2PO_4 + NaCl$. This is a particularly important detail of the buffer action of the blood, for it shows us how the phosphates of the corpuscles are rendered available for neutralizing acids added to the plasma, where there are practically no phosphates. Indeed the transference of acid through the corpuscular envelope indicates that the same sort of thing must go on with the other cells of the body, so that the plasma, itself rather poor in buffer substances, has all those of the body at its disposal.

THE MEASUREMENT OF THE RESERVE ALKALINITY

1. Titration Methods

There are several methods by which the reserve alkalinity of the blood may be measured. The simplest in theory consists in seeing how much standard acid must be added to a measured quantity of blood plasma in order to reach the neutral point as judged by change in tint of some indicator. The indicators employed (e. g., methyl orange) are such as change their tints at H-ion concentrations that are well to the acid side of neutrality (i. e., at a high C_H or low P_H). To bring the plasma to this point of neutrality the added alkali will need to neutralize, not only the bicarbonate of the plasma, but other acid-binding substances as well. This will give us a false impression of the acid-binding powers of the plasma, since, at the normal C_H of the blood, proteins do not absorb acids to anything like the extent they do at higher degrees of C_H. Another objection to the method is that the proteins interfere with the sensitiveness of the indicators.

The objections can be removed by determining the end point electrometrically or by indicators that change tint at about P_H7. The most practical way is to determine the change in C_H produced by adding a known volume of standard acid to blood plasma. The resulting change in C_H will then be greater the less the alkaline reserve. In the electrometric method irregularities that might be caused by variable amounts of carbonic acid in the blood to start with are best controlled by removing the CO_2 from the plasma after adding the standard acid. The procedure therefore consists in mixing 1 c.c. plasma with 2 c.c. N/50 HCl in a small separating funnel, which is then evacuated so as to remove the CO_2, after which the fluid is transferred to a hydrogen electrode and C_H measured (see page 29). In normal blood this should be $10^{-5.6}$ ($P_H5.6$). In acidosis, where there is a depleted alkaline reserve, the 2 c.c. of acid will cause a much greater change in C_H—in diabetic blood to below 5 or lower.

The technic involved in the above method is, however, too exacting for routine clinical work. For such purposes the colorimetric method of Levy and Rowntree may be employed.

THE METHOD OF LEVY AND ROWNTREE.[13]—A test tube made of hard ("nonsol") glass of about 20 c.c. capacity, containing about a gram of powdered *neutral* potassium oxalate, is filled with newly drawn blood, immediately stoppered and placed on ice. Quantities of 2 c.c. each of the blood are then placed in a series of seven small (nonsol) test tubes and allowed to stand for five to six minutes in order to permit a narrow

layer of plasma to separate on the surface (this prevents laking of the blood during the subsequent addition of acid or alkali). The blood in the first tube is used for the determination of the normal H-ion. In each of the next three tubes are added respectively 0.1, 0.2 and 0.3 c.c. N/50 HCl, and to the last three, similar quantities of N/50 NaOH. After inverting the tubes so as to mix the contents, the blood in each is transferred to celloidin sacs and the C_H determined according to the method described elsewhere (page 32).

The tubes are noted in which a change in tint from that of the normal blood is evident, and the results are expressed as the c.c. of N/50 HCl or NaOH which must be added to blood to change its C_H. Thus, the *alkali buffer* is the c.c. of N/50 NaOH which can be added to 2 c.c. of blood without change of C_H of the dialysate, and the *acid buffer* the c.c. of N/50 HCl.

The method suffers from the following drawbacks:

1. Very small quantities of acid and alkali are employed.
2. It is often difficult to tell just exactly when a slight difference in tint has been produced.
3. Even with the precautions described above, it is impossible to be sure that the amount of CO_2 in the different samples of blood is the same, which means, of course, that some bloods will, on this account alone, be able to bind more alkali than others.

THE METHOD OF VAN SLYKE.—A method based on somewhat the same principle, but which is more accurate because it meets the above objection, is that suggested by Van Slyke, Stillman and Cullen.[14] Plasma is freed of CO_2 by placing it in a vacuum, and is then mixed with an equal volume of N/50 HCl (or NaOH) and the C_H determined by the electric method (see page 29). In the case of normal blood, after such an addition of acid, a practically normal C_H will be found, whereas in the blood of cases of acidosis it will be very distinctly increased (i. e., P_H lower).

2. CO_2-combining Power

The above objections to the titration of blood plasma or dialysate with standard solutions of acids are removed if we measure the combining power of the blood alkali towards carbonic acid itself at normal blood reaction. This may be done either in blood immediately after its removal from the animal or in blood that has been first of all saturated outside the body with carbonic acid at a partial pressure equal to that existing in the body. Since for practical reasons venous blood must be used—in the clinic at least—the former of these methods suffers from the fault that varying amounts of carbonic acid will be added to the blood during its passage through the tissues, and the error thereby

incurred will become greatly aggravated if venous stasis has been produced in drawing the specimen for analysis. But the chief reason why this method has not been extensively employed, as pointed out by Van Slyke, is the technical difficulty of making the necessary analysis.

It is most satisfactory to collect venous blood after a period (one hour at least) of muscular rest (so that there is no excess of CO_2) and without venous stasis, and to centrifuge without permitting any considerable loss of carbonic acid. The latter precaution is necessary because there is a migration of acid radicles, e. g., HCl, from plasma into corpuscles when the CO_2 of the former is increased, and in the reverse direction when the CO_2 is decreased. If the CO_2 in the blood were not the same during centrifuging as it is in the body, the separate plasma would not contain the same amount of alkali—i. e., its reserve alkalinity would be altered. Although theoretically, therefore, centrifuging should be performed in

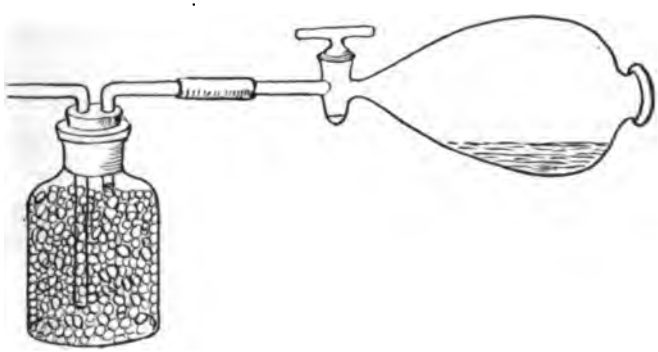


Fig. 10.—Diagram of apparatus for saturating blood or plasma with expired air. The glass beads in the bottle condense excess of moisture. The separating funnel, as soon as it has been filled with expired air, should be closed by a stopper and the stopcock turned off. It is then rotated so that the blood forms a film on its walls.

an atmosphere containing the same partial pressure of CO_2 as exists in the body (i. e., the alveolar air) (see page 344), this has been found impracticable for general use, and is unnecessary if loss of CO_2 from the specimen of blood is prevented by allowing it to flow into the syringe very slowly (without any suction). It is mixed in the syringe with powdered (neutral) potassium oxalate (enough to make a 1 per cent solution with the blood), and immediately delivered into a centrifuge tube under paraffin oil, which by floating on its surface serves to diminish free diffusion of CO_2 to the outside air (even though such oils dissolve more CO_2 than water). To mix the blood with the oxalate, the syringe should be moved backward and forward several times, but it must not be shaken.

After centrifuging, about 3 c.c. of plasma are removed and saturated with CO_2 at the same tension as in alveolar air (i. e., 5.5 per cent). This

is done by placing the plasma in a separating funnel of 300 c.c. capacity, laying the funnel on its side and displacing the air in it by alveolar air secured by quickly making as deep an inspiration as possible through the tube and bottle containing glass beads (Fig. 10). The glass beads remove excess of water vapor from the air. The funnel must be restoppered before the end of the expiration, so that no outside air enters. It is then rotated, for about two minutes, in such a way that the plasma forms a film on its walls. If it is necessary to postpone the saturating of the plasma, this should be pipetted off from the corpuscles and preserved in hard glass test tubes coated with paraffin. From ordinary glass

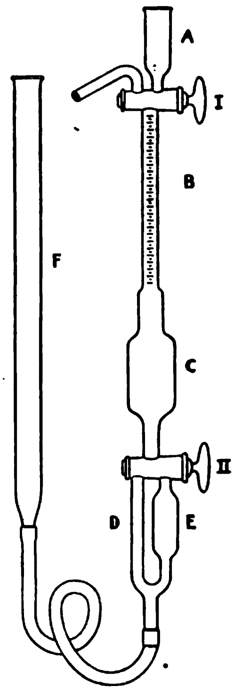


Fig. 11.—Van Slyke's apparatus for measuring the CO_2-combining power of blood in blood plasma. For description, see context.

enough alkali is soon dissolved out to vitiate the results. After saturation of the plasma with CO_2, the funnel is placed in the upright position and the plasma allowed to collect in the narrow portion, after which 1 c.c. is removed with an accurate pipette and analyzed for CO_2.

The analysis may be done by using either the Van Slyke or the Haldane-Barcroft apparatus. *The Van Slyke method is as follows:*

The apparatus is filled to the top of the graduated tube with mercury (Fig. 11) by raising the mercury reservoir *F*, care being taken that *D* and *E* are also filled. One c.c. of the CO_2-saturated plasma is then de-

livered into *A* (which has been rinsed out with CO_2-free ammonia water), and the stopcock *I* turned so that by cautiously lowering the level of the reservoir *F*, the plasma runs into *B* (but no trace of air). The same procedure is repeated with 1 c.c. water, so as to wash in all of the plasma, and finally 0.5 c.c of 5 per cent H_2SO_4 is sucked in, after which stopcock *I* is turned off. The reservoir *F* is then lowered sufficiently to allow all of the mercury, but none of the blood, to run out of *B* and *C*. A vacuum is thus produced in *B* and *C*.

As the level of the mercury falls in *B* and *C*, the plasma effervesces violently,* because it is exposed to a vacuum. To be certain that all traces of CO_2 have been dislodged from the solution, the apparatus is inverted several times. To ascertain how much CO_2 has been liberated, stopcock *II* is now turned so as to bring *C* and *E* into communication, and by cautiously lowering the reservoir the fluid in *C* is allowed to run into the bulb *E*. Stopcock *II* is thereafter turned so as to connect *C* and *D*, and the reservoir raised so that the mercury runs into *C* as far as the CO_2 that has collected in the burette will permit it to go. After bringing the level of the mercury in *F* to correspond to that in the burette, the graduation at which this stands is read. It gives the c.c. of CO_2 liberated from the plasma. Under the above conditions normal plasma binds about 75 per cent of its volume of CO_2; therefore, since the total capacity of the pipette is 50 c.c., the mercury should stand at 0.375 c.c. on the burette. For accurate measurement it is necessary to allow for the CO_2 that remains dissolved in the water, etc., as well as for barometric pressure and temperature. This is best done by the use of a table based on the known solubility of CO_2 under the various conditions obtaining, which is given in Van Slyke's paper.[12]

The Haldane-Barcroft apparatus that is most suitable for the above analysis is shown in Fig. 136, page 382.† One c.c. of CO_2-free ammonia water is placed in the bottle and the 1 c.c. of plasma delivered beneath it.

*This may be prevented by adding a small drop of caprylic alcohol.

†This form of Haldane-Barcroft apparatus is not quite the same as the differential manometer that is used for measurement of the O_2-combining power of hemoglobin (page 382). In the form used for the present purpose, a side tube at the bend of the U-tube is connected with a small rubber bag, which can be compressed by a screw. When the gas is evolved in the bottle, it presses down the fluid in the proximal limb of the manometer correspondingly and raises that in the distal limb. Since the calculation of the amount of gas evolved depends on finding the *pressure* produced without any change in volume, it is necessary after the gas has been evolved to compress the rubber bag until the meniscus of fluid in the proximal limb of the manometer is brought back to its original level. The height at which the fluid stands in the distal limb then obviously corresponds to the pressure created by the evolved gas.

The equation for determining the amount of gas evolved depends on the gas law, which states that the pressure of a gas is inversely proportional to its volume (page 336). Suppose that the volume of gas evolved was equal to the volume of the bottle, then, since the volume has been kept constant, the pressure would be doubled—that is, the fluid in the distal limb would equal that of 1 atmosphere, or 10,400 mm. of water or 10,000 of clove oil, which is the fluid actually used to fill the manometer. Any other observed pressure would therefore correspond to the volume of evolved gas according to the equation,

$$V = \frac{\text{vol. of bottle (and tubing to meniscus)}}{\text{10,000 (when clove oil is used)}}$$

In using the apparatus in the above manner, only one of the bottles is employed, and the tartaric acid is added from a pocket in the stopper by a simple manipulation.

The bottle is then connected with the manometer with the precautions described elsewhere in this volume. When temperature conditions have been allowed for, saturated tartaric acid is mixed with the plasma solution and the gas evolved measured by the displacement of the fluid in the manometer. The apparatus may also be used with blood in place of plasma. In this case, however, it is necessary that the oxygen be removed before adding the tartaric acid. This precaution is necessary, since acid can dislodge some of the O_2 from hemoglobin. The blood is therefore first of all laked with ammonia containing some saponin, then shaken with 0.25 c.c. saturated potassium ferricyanide solution, and finally with the saturated acid solution. If blood is used, the estimations must be made on strictly fresh blood, since on standing the CO_2-combining power greatly deteriorates.

3. Indirect Methods

There are several other methods by which the alkaline reserve may be measured. These include:

1. **Determination of the Tension of CO_2 in Alveolar Air** (page 344).—Since this method is employed more particularly in investigating the hormone control of the respiratory center, we shall defer a description of it until later. The alveolar CO_2 tension corresponds to the CO_2 tension in arterial blood and this is proportional to the alkaline reserve as determined by Van Slyke's method as is proved by the fact that the ratio, $\frac{\text{plasma } CO_2}{\text{alveolar } CO_2 \text{ tension}}$, is satisfactorily constant.

2. **The Measurement of the Acid Excretion by the Kidney.**—As might be expected, the acid-base equilibrium of the body may also be gauged by *measurement of the acid excretion of the urine,* in which the acids are contained partly in combination with ammonia or a fixed base, and partly in a free state. We shall first of all consider the methods of acid excretion and then examine the evidence showing that the total acid excretion is proportional to the alkaline reserve as measured by the above described methods.

Excretion of Acid in Combination with Ammonia.—The production of ammonia is essentially an endogenous process, and when excessive quantities of acid make their appearance in the organism, the fixed alkali may not be sufficient to neutralize it all, so that ammonia, derived from the breakdown of amino acids (page 616), instead of being converted into urea is employed to neutralize the excess of acid. Most workers have in this way explained the very large ammonia excretion that has long been known to occur in such conditions as diabetic acidosis. Some recent workers are, however, inclined to question the significance of

ammonia in this connection, believing that the increased ammonia excretion is, like the acetone bodies themselves, a product of perverted metabolism. Be this as it may, it is no doubt true that ammonia is used for neutralizing acid in disease, although it may not be an important factor in the maintenance of neutrality under normal conditions. It is a factor of safety, in that it helps to care for an increase in acid when the normal mechanism of the body is overtaxed.

Excretion of Phosphates.—The more permanent control of neutrality depends on the excretion of phosphates by the kidney. The principle governing this process is exactly the same as that already discussed in connection with carbonic acid. In the one case it is the volatile acid CO_2, and in the other, the fixed phosphoric acid that is concerned in the reaction. The ratio between the acid salts of phosphoric acid, MH_2PO_4, and the alkaline salts, M_2HPO_4, in blood is approximately 1 to 5, but in the urine this ratio varies according to the amount of H ion that must be eliminated from the blood. In other words, a definite amount of phosphoric acid is enabled to carry variable amounts of H ion out of the body by causing the amount of alkali excreted in combination with it to become altered. For example, in the form of MH_2PO_4 a given amount of PO_4 obviously carries out more H ion than when it is excreted as M_2HPO_4. The adjustment between these two salts is a function of the kidney. We may accordingly measure the amount of alkali retained by the organism by finding how much standardized alkali must be added to a given quantity of urine until the reaction of the blood is obtained. Since the latter value is constant, the titration can be done simply by titrating the urine with an indicator such as sulphonephenolphthalein, which changes tint at about P_H of blood.

A more serviceable indicator to use, however, is *phenolphthalein*, because its end point is such that when human urine just reacts neutral to it—that is, when the titrable acid approaches zero—the CO_2-absorbing power of the plasma is at its maximum of 80 vols. per cent and the ammonia excretion by the urine is zero (Van Slyke). It is advantageous, therefore, to use this indicator, because it happens to have its turning point situated for a reaction which is well to the alkaline side of neutrality, and which is reached in urine when the blood is at its maximal acid-combining power and no ammonia is being used for neutralization purposes. As the CO_2-combining power of the blood decreases, there should, therefore, be a proportionate increase in ammonia and in the titrable acidity of the urine.

Although a general parallelism exists between these values in cases of diabetes, etc., there is no strict proportionality. The expedient has therefore been tried of comparing the alkaline reserve of the blood with

the excretion rate of acid as determined by an application of Ambard's equation for chlorides and urea, and with curiously satisfactory results (Fitz and Van Slyke). This equation is:

Blood concentration $= \text{constant} \times \sqrt{\frac{D}{W}}\sqrt{C}$; where D is the excretion rate, W the body weight, and C the concentration of excretory product in the urine. For the present purpose D is therefore the number of c.c. of N/10 alkali (or acid) required to bring the urine to the neutral point of phenolphthalein *plus* the NH_3 expressed as c.c. of an N/10 solution, for the twenty-four hours, and C is c.c. of N/10 alkali and of N/10 NH_3 per liter of urine. If we assume that the acid accumulation in the blood is proportional to the fall of the plasma CO_2 figure below the maximal figure of 80, the above equation becomes:

$$\text{Retained acid} = 80 - \text{plasma } CO_2 = \text{constant} \times \sqrt{\frac{D}{W}}\sqrt{C}.$$

For practical purposes it is best to make the necessary analysis on a sample of urine collected over a period of one to four hours, and to collect the blood for determination of its reserve alkalinity in the middle of this period. The twenty-four-hour rate of excretion is then computed (D) from the analysis.

The value calculated by the above equation has been found to agree with that of the CO_2-combining power of the plasma to within 10 volumes per cent, except when bicarbonate is being taken by the person, when the blood bicarbonate is much higher than indicated by the urine.

3. **Determination of Alkali Retention.**—Another valuable criterion of the alkaline reserve is the amount of alkali required to change the reaction of the urine. In health the C_H of the urine varies from 0.000,016 *N* ($P_H = 4.8$) to about 0.000,000,035 *N* ($P_H = 7.46$) with a mean of about 0.000,001 *N* ($P_H = 6$). These extremes are rarely overstepped in disease, but frequently the *average* is considerably different. In cardio-renal disease, for example, the mean acidity may be approximately 0.000,005 *N* ($P_H = 5.3$), or five times the normal value. A certain degree of acidosis is therefore common enough in this condition—a fact which has indicated the advisability of administering sodium bicarbonate. It has been found that 5 grams or less of soda, given by mouth to a normal person, causes a distinct diminution in the C_H of the urine, whereas in pathologic cases it may be necessary to give more than 100 grams before a similar effect is observed (L. J. Henderson and Palmer[15] and Sellards[16]).

For this very large holding back of alkali, the organism and not the kidney is responsible. This is indicated by the fact that, when the administration of alkali is discontinued, the acidity of the urine soon

regains its old level, although now if a smaller dose of alkali is given, the C_H of the urine will immediately be lowered. These facts indicate that for the moderate degrees of acidosis common in chronic disease, the properly controlled administration of soda is very probably a most advantageous treatment.

CHAPTER VII

COLLOIDS

Substances which can be obtained in the crystalline state and which, when in solution, are capable of readily diffusing through membranes, are designated as crystalloids, and are to be distinguished from another, larger group of substances not having these characteristics or having them only in very minor degree—the colloids. In every field of chemistry the properties of colloids have been studied extensively during recent years, but in no field more than in that which covers the chemistry of biological fluids and tissues, into whose composition colloids enter much more extensively than crystalloids. The subject of colloidal chemistry has indeed become so extensive that an attempt to do more than indicate some of the most important characteristics of colloids would take us far beyond the limitations of this book. The far-reaching applications of the subject in physiology and medicine are only beginning to be realized.

The term "colloid," or "colloidal," does not refer to a class of chemical substances, but rather to a state of matter which is quite independent of the chemical composition of the substance. We are familiar with more colloids in the organic than in the inorganic world, yet they are plentiful in both, and the same substance may at one time be colloidal and at another noncolloidal. Indeed, under appropriate conditions probably all substances may assume the colloidal state—not solids and liquids alone, but gases as well. It is mainly with liquids, however, that we are concerned in biochemistry.

CHARACTERISTIC PROPERTIES

The distinction between molecular* and colloidal solutions is a relative one. Suppose, for example, that we take a piece of gold in water and divide it up into smaller and smaller parts. At a certain stage, the particles will be so fine that they will remain in suspension and be invisible by ordinary means. They are then said to be in the colloidal state. If we divide them further until they become molecules of gold, a molecular solution will be obtained. In the colloidal state, there are

*Molecular solutions include those of nonelectrolytes, such as sugar, and electrolytes, such as inorganic salts.

two distinct phases in the solution, one solid and the other liquid, and between the two, because of the great subdivision of the original particle, is an enormous surface of contact. The solution is *heterogeneous*, and at the interface between the two "phases" the physical forces which depend on surface—e. g., surface tension (see page 64)—are enormously developed, and are responsible for the peculiar properties of colloidal solutions as compared with those of molecular solutions, which may, therefore, be styled *homogeneous*. The solutions of crystalline substances which we have hitherto been concerned with, are homogeneous.

Between these two groups of solutions is an intermediate one—namely, *suspensions* (as suspensions of quartz or carbon, or oil emulsions). Besides being turbid in transmitted light, the solutions may be seen by means of the ultramicroscope to contain particles. These can be separated by filtration from the fluid they are suspended in, except in the case of many emulsions in which the particles can squeeze their way through the filter pores by changing their shape. On standing or being centrifuged suspensions may also separate into their constituents, although this can be greatly hindered by the addition of a suspending substance such as gelatin or certain bodies having a so-called protective action (as peptone, proteose, etc.).

True Colloidal Solutions

1. **The Solution Is More or Less Turbid.**—Frequently this can be recognized by holding the solution in a thin-walled glass vessel against a dark background, but the turbidity may be so slight that it requires for its detection the use of the Tyndall phenomenon. This is familiar to all in the effect of a beam of sunlight let in through a small aperture into an otherwise darkened room. In the course of the beam suspended dust particles, which are invisible in an equally illuminated room, become visible, and thus render very distinct the pathway of the beam. If a colloidal solution contained in a glass vessel, preferably with parallel sides, is held in the course of such a beam, *the Tyndall phenomenon* will be seen in the liquid, which is not the case with molecular solutions. Focused artificial light may be employed for intensifying the effect. The light that is sent out at right angles to the beam is plane-polarized, which means that the particles reflecting the light must be smaller than the mean wave length of the light forming the beam. It should be understood that the individual particles themselves may not be rendered visible to the naked eye by the beam, although in such cases they can often be seen by using intense illumination and a dark-field (ultramicroscope) combined with suitable magnification (Fig. 12).

2. **Colloids Do Not Readily Diffuse.**—To demonstrate this, test tubes

are half filled with a 5 per cent solution of pure gelatin or a 1 per cent solution of pure agar, and, after the jelly is set, the solution under examination is poured on the surface; or, when it is of high specific gravity, the tube of gelatin, etc., is placed mouth downwards in the solution. In the case of colloidal solutions very little if any diffusion into the gelatin or agar will occur, even after several days; whereas true molecular solutions will diffuse for a considerable distance. When colored solutions are used, the diffusion can readily be recognized by visual inspection (see Fig. 13), but when they are colorless, the presence or absence of diffusion must be determined by removing the column of gelatin or agar and dividing it into slices of equal size, which are then examined chemically for the substance in question.

A further test is afforded by the failure of colloids to diffuse through membranes (dialysis). This was the method originally used by Thomas Graham to distinguish between molecular and colloidal solutions. The solution under examination is placed in *a dialyzer,* which is then immersed in a wide vessel containing the pure solvent. The older forms

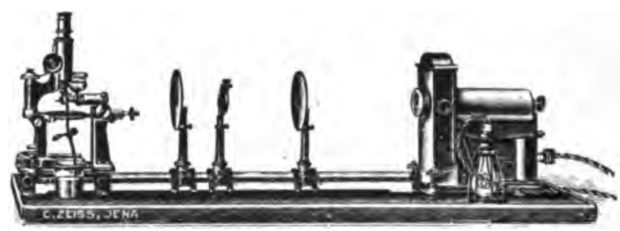


Fig. 12.—Ultramicroscope (slit type) for the examination of colloidal solutions. The arrangement of diaphragms, etc., in this form removes the absorptive effects of the surfaces of the glass vessel or slide used to contain the colloidal solutions.

of dialyzer consisted in general of a bell-shaped glass vessel closed below with parchment paper, but more recently so-called diffusion sacs have been adopted. These consist of pig or fish bladders or of collodion sacs. The latter are made by placing some collodion dissolved in ether in a test tube, which is then tilted so that the collodion runs out except for a thin layer which remains adherent to the walls. When the collodion has set, the sac can be removed after loosening it by allowing a little water to flow between the sac and the walls of the test tube. The sac containing the colloidal solution is then suspended in water or some of the solvent used in preparing the colloidal solution, care being taken that the menisci of the fluids inside and outside of the sac stand at the same level. Sometimes, especially when collodion sacs are used, some colloid may at first diffuse through, but if the outer fluid (the dialysate) is renewed and the dialysis allowed to proceed, this ceases.

When a fluid solution exhibits both of the above properties (i. e., the Tyndall phenomenon and indiffusibility), there can be no doubt as to its being in a true colloidal state, but there are substances, such as congo red or protein solutions of certain strengths, which may exhibit a very slight diffusibility in a dialyzer but not show the Tyndall phenomenon. Substances of this group constitute transitional types between molecular and colloidal solutions, and to determine their true nature it is neces-

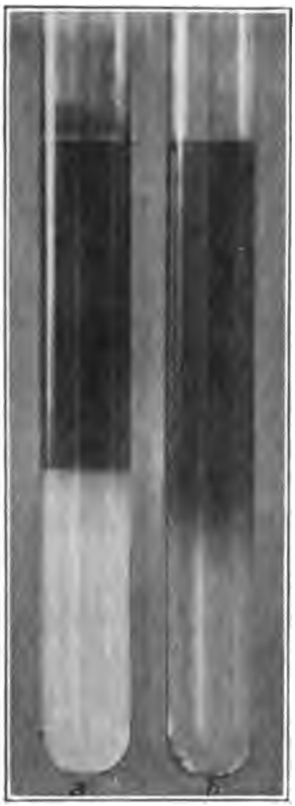


Fig. 13.—To show diffusion into gelatin of a crystalloid stain in *b* and the nondiffusion of a colloid stain in *a*. (From W. Ostwald.)

sary to employ refined methods such as those of ultramicroscopy, ultrafiltration, etc., which can not be described here.

3. The Size of Colloidal Particles.—It will be apparent that the essential property upon which the above-mentioned phenomena depend is the size of the particle. Particles which can still be seen under the microscope are called *microns*. They have been computed to have a dimension of 0.1 μ (0.001 mm.) or more, and they form suspensions. Particles which are invisible microscopically under the ordinary conditions of illumina-

tion, but are still visible when the ultramicroscopic illumination is used, are called *submicrons.* They have a dimension between 0.1 μ and 1 $\mu\mu$ (0.000,001 mm.),* and they constitute the colloids. Particles smaller than 1 $\mu\mu$ are called amicrons, this term being used to include the molecules and ions present in molecular solutions. (The amicron of hydrogen is, for example, computed to be 0.067 to 0.159 $\mu\mu$, and that of water vapor, 0.113 $\mu\mu$.) This classification of dissolved substances according to the size of the particles and molecules shows the relationship of one

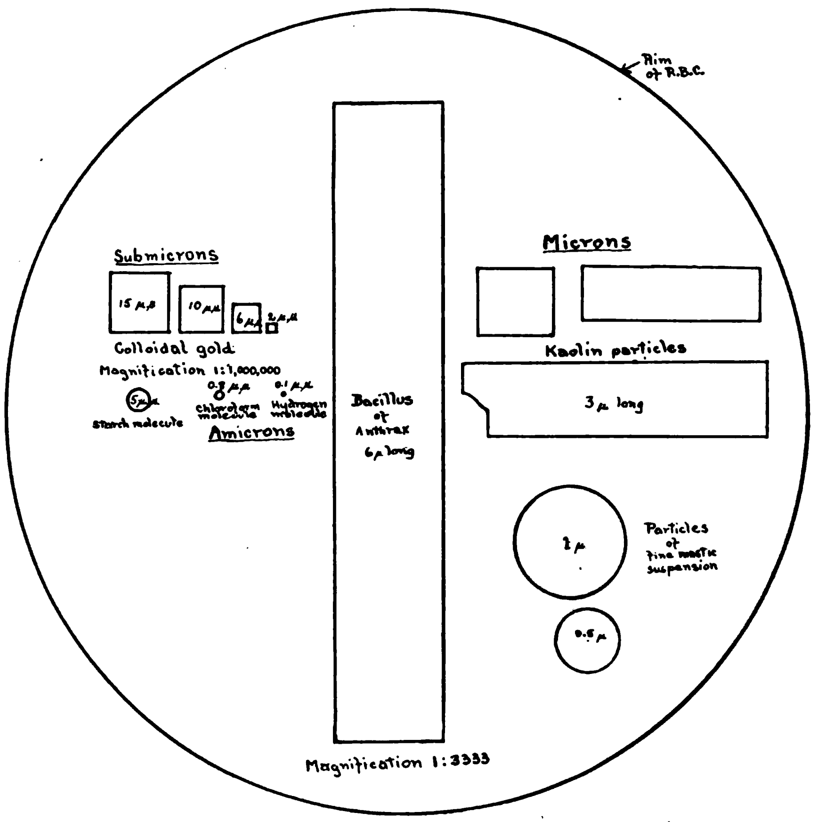


Fig. 14.—Diagram from W. Ostwald showing the relative size of various particles and colloidal dispersoids compared with a red blood corpuscle and an anthrax bacillus.

class of substances to others. An idea of the relative sizes of colloidal particles and molecules in comparison with such familiar objects as a blood corpuscle and an anthrax bacillus is given in Fig. 14. The fluid in which the "particle" is suspended is called the *dispersion medium,* or *external phase,* and the particle itself the *dispersoid,* or *internal phase.*

It is the enormous development of surface which determines the dif-

*μ = 0.001 mm., and $\mu\mu$ = 0.000,001 mm.

ference in the properties of a colloidal solution from those of a suspension of the same substance. Thus, the difference between a colloidal solution of platinum (prepared by allowing an electric arc to form between platinum electrodes in water) and pieces of platinum in water depends on the fact that the surface of the platinum in the former case has been increased many million times. When the subdivision becomes still greater and the particles gain the size of molecules, the phenomena due to surface development become suppressed and those due to concentration in unit volume become accentuated. The properties dependent on osmotic pressure, diffusibility, etc., are exhibited by all dispersoids, whether ions, molecules or particles, but some of these properties are much more pronounced when the dispersoids are of large dimensions, and others when they are small. In other words, the phenomena due to surface, such as those of surface tension (see page 64), become apparent only when the dispersoids have the properties of matter in mass; when the dispersoids become molecular in size, they manifest the properties characteristic of true solutions.

4. Electric Properties of Colloids.—Most colloids carry a charge, which may be either positive or negative toward the dispersion medium. Both crystalloids and colloids therefore carry electric charges; in the former case, however, the charge does not reveal itself until the molecules in solution have become dissociated, when each ion carries a charge of opposite sign (see page 16), whereas in the case of colloids, each colloid particle usually carries a charge which is always of one sign, either positive or negative. Colloids may therefore be grouped into positive and negative, according to the charges which they carry, and there is a third group in which the charge may be either positive or negative according to the nature of the dispersion medium.

A colloid not carrying a charge to begin with can be caused to assume one by the action of electrolytes, for the electrical properties of colloids, as well as those of inert powders suspended in water, are readily influenced by the charges present in the ions of the dispersion medium. The $H^{\cdot}$ and OH' ions are especially liable to exert this influence. The particles of inert powders in suspensions (kaolin, sulphur, etc.) carry a positive charge when the water in which they are suspended is acidified, and a negative charge when it is made alkaline. In general, it may be said that suspensions of most powders and of insoluble organic acids in water (e. g., charcoal, cellulose, kaolin, caseinogen, mastic, free acid of congo red, etc.) are electro-negative. Of true colloids ferric hydroxide (ferrum dialysatum) and serum globulin are positive in acid solutions; arsenious sulphide and serum globulin are negative in alkaline solution, and serum globulin in neutral solutions has no charge.

To ascertain the nature of the charge various methods may be employed, of which the following are important:

1. The method of *electrophoresis*. The colloid solution is placed in a U-tube, each side of which carries a platinum electrode dipping into the solution. After a strong continuous electric current has been allowed to pass for some time through the solution, it will be found that the colloid collects at the anode (where the current enters) when it is a negative colloid (since unlike electric charges attract each other), and at the cathode when it is positive. In the case of colored solutions, the migration can be readily seen, but otherwise it may be necessary to analyze the solution at the two poles.

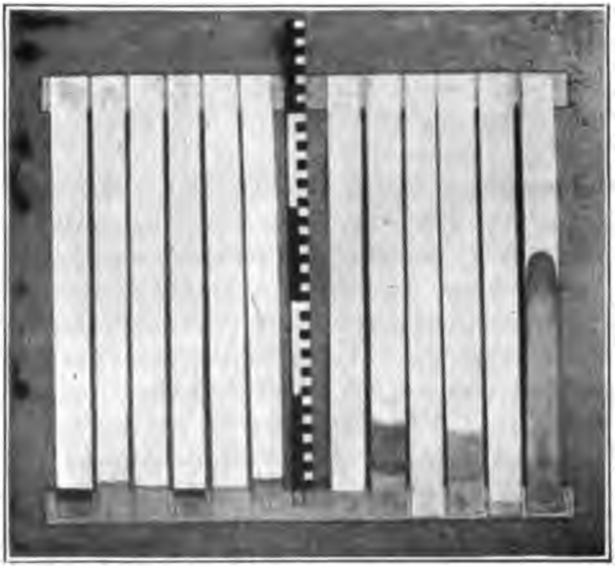


Fig. 15.—Capillary analysis of colloids. Strips of filter paper, after being suspended with the lower ends dipping into colloidal solutions. Those on the right hand were positive colloids, which did not rise in the strips, but formed a sharp line of demarcation at the lower end on account of precipitation. Those on the left hand were negative colloids. (From W. Ostwald.)

2. The method of *capillary analysis*. For this purpose a long strip of filter paper is arranged vertically over the solution, with its lower end dipping into it. In the case of negative colloids the colloid, as well as the dispersion medium, rises uniformly on the strip of paper (it may be to a height of 20 cm.); whereas with positive colloids the dispersion medium alone rises, the colloid itself doing so only to a very slight extent, but becoming so highly concentrated at the interface between the solution and the paper that it coagulates on the end of the strip of paper, where it forms a sharp line of demarcation (Fig. 15).

3. The method of *mutual precipitation* of colloids. When a positive

and a negative colloid are mixed in such proportions that the electric charges are neutralized, precipitation usually occurs. When it does so, we can tell the nature of the electric charge of an unknown colloid by its behavior when a colloid of known electric sign is added to it. For example, if ferric hydroxide (positive) causes a precipitate to form when it is added to an unknown colloidal solution, the electric charge of the latter must be negative; if it does not precipitate with ferric hydroxide, but does so with arsenious sulphide (negative), it must be positive.

5. Brownian Movement.—Like the particles in fine mechanical suspensions, those of colloidal solutions, especially when examined ultramicroscopically, exhibit the so-called Brownian movements, which have been described as "dancing, hopping and skipping." These movements occur in straight lines, which are suddenly changed in direction and are quite independent of external sources of energy, such as change in temperature (although they become quicker as the temperature of the solution is raised), earth vibrations, chemical changes, or the electric charge of the colloid. The movements become more rapid the smaller the particles, and they become sluggish as the viscosity of the solution increases. Addition of electrolytes decreases the movement by causing the particles to clump together. The density and viscosity of the dispersion medium, the electric charge of the dispersoid and the presence of Brownian movements, are the forces which operate together to prevent sedimentation of the particles in a colloidal solution.

6. Osmotic Pressure.—As one of the distinguishing properties of colloids we have seen that their diffusibility, as into gelatin or agar jellies, is extremely slow when compared with that of a molecular solution. This does not mean, however, that colloids are possessed of no power of diffusibility if left long enough. Indeed the existence of the Brownian movement indicates that such diffusion *must* occur, and therefore it should be possible, by the application of the same principles as those which govern molecular solutions (e. g., by using a semipermeable membrane), to measure the osmotic pressure.

Many studies of the osmotic properties of colloidal solutions have been undertaken, especially by those who are interested in the possibility that the colloids of blood serum (serum albumin and globulin) may create an osmotic pressure. If this should prove to be the case, it would be necessary for the osmotic pressure to be overcome by mechanical pressure such as that supplied by the heart (i. e., the blood pressure) in the various physiologic processes of filtration and diffusion taking place through cell membranes (as in the formation of urine in the kidney).

For measuring the osmotic pressure of colloids, osmometers similar

to those already described (page 4) can be employed. Most of the recent work has been done either with collodion sacs, or with unglazed clay cups impregnated with some gel, such as silica or gelatin. When such an osmometer, filled with some colloidal solution (like a solution of pure albumin) and provided with a vertical glass tube, is placed in an outer vessel containing water, the fluid will be seen to rise in the vertical tube, the height to which it rises being proportional to the osmotic pressure.

But the observed pressure does not necessarily give us the osmotic pressure of the pure colloid, for to this, even when highly purified, there is almost certain to be attached a considerable amount of inorganic salt, which may be responsible for the osmosis. It has indeed been maintained by some observers that electrolytes form an integral part of certain colloids, being bound to them perhaps by adsorption (see page 65), and that they are essential to the maintenance of the colloidal state. In any case, since electrolytes are always present, the osmotic pressure of the pure colloid can be measured only when means are taken to discount their influence. Several devices have been used, of which the following may be mentioned:

1. Addition to the fluid outside the osmometer of a percentage of salt equal to that found by chemical analysis to be present in the colloid. (This method is untrustworthy.)

2. The use of a limited quantity of fluid on the outside of the osmometer so that equality of saline content soon becomes established, by diffusion, in the fluids on the two sides of the membrane.

3. The use of a membrane which is permeable to electrolytes but not to colloids.

Even when the greatest care is taken in its measurement, the osmotic pressure of a given colloid has been found to vary considerably not only according to the method used in its preparation, but also according to the amount of mechanical agitation (shaking, stirring, etc.) to which the colloid solution has been subjected. Regarding the influence of the method of preparation, it was found in one series of experiments that albumin that had been repeatedly washed (but still contained considerable ash) gave no osmotic pressure, whereas another preparation that had been purified by crystallization twice (and contained much less ash) had a pressure of 3.38 mm. Hg. According to these results the ash content of the colloid is not fundamentally responsible for its osmotic pressure. As to the influence of mechanical agitation, the osmotic pressure of a gelatin solution is increased by shaking, while that of a solution of egg albumin is decreased.

The property upon which the osmotic pressure depends is undoubtedly

the state of dispersion of the colloid particles, and until we know all of the factors which may influence this, measurements of osmotic pressures of colloids can scarcely be of very much value. Nevertheless, that this property has some physiologic bearing is clear from the effect which colloids have in restoring the blood pressure after hemorrhage (page 141).

Further evidence that the osmotic pressure of colloids has not the significance that it has in the case of molecular solutions is furnished by the fact that the osmotic pressure is only approximately proportional to the concentration of the solution; it may either increase or decrease relatively to the strength of the solution. Temperature also has quite a different influence on the osmotic pressure of colloids from that which it has on the osmotic pressure of molecular solutions, and it frequently has an influence which persists after the solution is brought back to its original level.

The influence of added substances on the osmotic pressure of colloidal solutions is of considerable interest to the biologist, for, whereas in the case of molecular solutions this is purely additive, in the case of colloids the added substance may at one time cause the osmotic pressure to increase, at another, to decrease. It has been found that the osmotic pressure of gelatin solutions at first decreases, then rapidly increases as the H-ion concentration is raised. The addition of alkali increases the osmotic pressure until a maximum is reached, beyond which it begins to fall. Both acids and alkalies lessen the osmotic pressure of egg albumin. Electrolytes always decrease the osmotic pressure of gelatin and albumin solutions, and the degree to which they exert this influence depends on the nature of the cation and anion composing the electrolyte. In the order of their depressing influence the *cations* arrange themselves:

Heavy metals > alkaline earths > alkalies;

and the *anions:*

$$SO_4 > Cl > NO_2 > Br > I > CNS.$$

The influence of a given electrolyte varies extraordinarily with the reaction of the colloid, a fact which must be carefully regarded in all work in this field.

CHAPTER VIII

COLLOIDS (Cont'd)

SUSPENSOIDS AND EMULSOIDS

According to whether colloids form solutions that are more or less viscid than the suspension medium, they are divided into *emulsoids* and *suspensoids*. Examples of the former class are silicates and gelatin, and of the latter, dialyzed iron and arsenious sulphide. The following characteristics are used to distinguish between suspensoids and emulsoids:

1. Measuring the time it takes, at a standard temperature, for a given volume of the fluid to flow out of a standard pipette (10 c.c.) shows the *viscosity* to be, roughly, inversely proportional to the time of outflow. In the case of suspensoids the viscosity is no different from that of the dispersion medium alone, and does not vary much when the solution is cooled. The viscosity of emulsoids even in very dilute solutions is, on the other hand, considerably greater than that of the dispersion medium itself, and it becomes greatly increased by cooling.

2. Suspensoids are much more readily coagulated by the addition of electrolytes than emulsoids. This is particularly true when water is the dispersion medium (so-called hydrosols), and when electrolytes having a polyvalent ion (such as Al or Mg.) are employed. Thus, practically all suspensoids are coagulated in the presence of 1 per cent of alum, which has no influence on emulsoids. We shall return to this phase of our subject later on.

The division of colloids into emulsoids and suspensoids is more or less arbitrary, since one class may be changed into the other, the determining factor being the water content of the dispersoid. The water content of suspensoids is low (lyophobe), while that of emulsoids is high. By changing the relative amounts of water and solid of which a colloidal solution is composed, the nature of the dispersoid may be changed. If the water is diminished, the dispersoid behaves as a suspensoid and becomes readily precipitated. The practical importance of this fact is that it explains *the salting out of proteins*—a process extensively used in their separation. Ordinarily these behave as emulsoids, but the addition of salt raises the osmotic pressure of the dispersion medium, and thus attracts water from the dispersoids, with the result that they come

to behave as suspensoids, and are accordingly precipitated by the electrolytes.

Another property of emulsoids of biological importance is the protection which they can afford against the precipitating influence of electrolytes on suspensoids. If a colloidal solution of gold is mixed with a trace of gelatin, the subsequent addition of salts will be found to produce no precipitation. The explanation of this is that the emulsoid becomes distributed as a film on the suspensoid particles, thus practically converting them into emulsoids.

Gelatinization

One of the best known properties of emulsoids is that of gelatinization, which has an interesting bearing on many problems of biology. After the gel has set, an enormous pressure is required to squeeze out any water from it, indicating that the water no longer forms the continuous phase but must be enclosed in vesicles formed of more solid material.

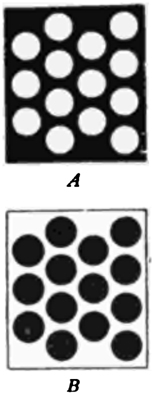


Fig. 16.

As a gelatin solution cools, the gel at first forms a polarized cone of light, but the very fine particles which are responsible for this effect soon increase in number and size so that they obstruct one another in their Brownian movements and adhere, giving an appearance of fine felt-like threads throughout the solution. A sort of impervious sponge work of the more solid phase is therefore formed, the more fluid phase being inclosed in the meshes.

If, as in the accompanying diagram, the dispersion medium is represented by white and the dispersoid in black, the relationship between the two in a suspensoid is as in *A*, and that in a gel as in *B*. To express any of the dispersion medium in *B*, it will require a pressure sufficient to

cause the more fluid phase to penetrate the more solid. If the gel is treated with reagents like formaldehyde, the liquid can be readily pressed out. This occurs during fixation for histological purposes.

Imbibition

Closely related to gel formation is the process of imbibition—the power of taking up large quantities of water without actually forming liquid solutions. Besides gelatin the dried tissues of plants and animals exhibit the phenomenon, and it is undoubtedly of importance in many physiologic processes such as growth and the passage of water into cells, etc. The materials present as vacuoles in plant cells attract water from the environment of the cell by imbibition, and thus exert on the cell wall a pressure which, acting along with the osmotic pressure, maintains the turgor of the cell. The initial growth of pollen is also dependent upon imbibition, and important observations on this process under varying conditions, are likely to furnish us with useful information concerning the significance of imbibition in connection with growth of cells in general.

By measuring the rate of increase in length of long, narrow strips of gelatin placed in Petri dishes containing solutions of varying composition, the factors that influence the imbibition process can be quantitatively investigated. Working in this way, F. H. Lloyd[17] has found that for all acids there is a certain concentration (about N/320 H_2SO_4) which induces a maximum rate of swelling, and another, much weaker (N/2800 H_2SO_4), in which the rate of swelling is even less than in pure water. In higher concentrations of acid than N/320, the gelatin at first swells very quickly, but the rate slows off so that it soon comes to be less than that with intermediate concentrations. Regarding alkalies, at high concentrations the effect is similar to that of acids. Salts alone seem to repress the swelling below that of water. It should be pointed out that the concentrations of acid and alkali in the above observations are much greater than those that could occur in the animal body. The experiments recall the attempts made some years ago by Martin Fischer to explain edema as due to excessive imbibition of water by the proteins of the tissues because of increased acidity of the blood and tissue fluids. That imbibition might possibly play some role in such processes is not denied, but Fischer disregards entirely the now well-established facts that hydrogen-ion concentration is one of the most constant properties of the blood, that very low concentrations of acid may diminish rather than increase imbibition, and that it is manifested only in the absence of inorganic salts.* Moreover, the fluid in edema can often

*Determinations of the hydrogen-ion concentration of the blood recently published from Fischer's laboratory do not inspire confidence.

be drained off by hollow needles, and it passes by gravity from one part of the blood to another, neither of which processes would be possible if imbibition were the essential factor concerned. If further evidence against this hypothesis should be demanded, it might be found in the utter failure of the therapeutic measures—alkali administration—that are recommended to combat the edema.

Action of Electrolytes on Colloids (apart from their effect on osmotic pressure).—It has been stated above that the charge which a colloidal particle assumes may be neutralized by a charge of opposite sign carried by an ion present in the dispersion medium. The neutralization of the electric charge causes coagulation of the suspensoids but not of the emulsoids. Of the positive and negative ions into which the electrolytes dissociate, the one producing the coagulation is that which is opposite in sign to the electric charge of the colloidal particle.

A quantity of electrolyte which is capable of producing complete precipitation when added all at once to suspensoids will be ineffective when added in small quantities at a time. This phenomenon, which is also known to be exhibited when toxins and antitoxins are mixed together, is probably owing to the fact that precipitation depends on inequality and irregular distribution of electric charges, a condition which becomes established when the electrolyte is suddenly added, but not so when it is gradually added. The particles in the latter case become, as it were, acclimated to the electric charges introduced by the addition of the electrolyte.

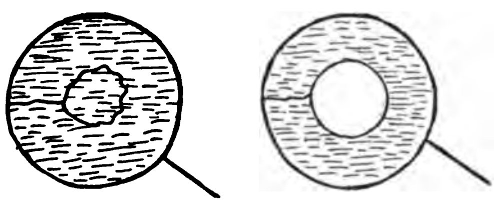
Proteins as Colloids.—The most prominent colloids in the field of biochemistry are the proteins. On account of complexity of structure, however, certain factors intervene which render the investigation of their behavior very difficult. As we shall see later, proteins are made up of combinations of amino acids, each of which contains basic (NH_2) and acid groups (COOH). The various amino acids are linked together in protein by the COOH of one uniting with the NH_2 of another, with the elimination of water—thus, CO $\boxed{OH + H}$ HN—but some NH_2 and COOH groups are left uncombined. According to the relative number of these uncombined radicles, the protein (or polypeptid, page 601) will exhibit faintly acid or basic or neutral properties. With acids, for example, a salt will be formed by union with the NH_2 groups, which will dissociate into the anion of the acid and a large organic cation; whereas with alkalies union will occur with the COOH group, and the salt on dissociating will form a small cation of the metal of the salt and a large complex anion. We may therefore obtain the protein with either a positive or a negative electric charge by altering the chemical nature of

the fluid in which it is dissolved, so that the reaction towards other colloids and towards electrolytes will vary.

One feature of proteins of importance in this connection is that known as the *isoelectric point,* at which the protein exists with a maximum of electrically neutral molecules. This point is reached by adding acid to a protein solution. The acid represses the dissociation of the protein acting as an acid, and therefore diminishes the number of free hydrogen ions; and at the same time it combines with the NH_2 groups and neutralizes the basic characteristics. The alteration in electric charge thus induced alters the water-absorbing powers of the protein and therefore all of the properties which we have seen to be associated therewith (page 63).

SURFACE TENSION

Before we consider a very important property of colloids known as adsorption, by means of which they are able to perform many reactions that do not conform with the laws of mass action, it will be well to



A. *B.*

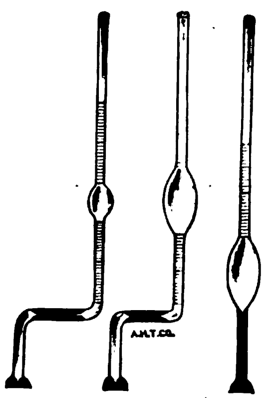
Fig. 17.—Diagram to illustrate surface tension. The rings *A* and *B* inclose soap films in which a very fine loop of silk is suspended. In *A* it is loose but in *B*, where the film inclosed in the loop has been broken, it is drawn into a circle by the tension of the soap film. (From Bayliss.)

say a few words concerning the physical phenomenon upon which this depends—namely, surface tension. The creation of this force is due to the fact that, whereas the molecules within a liquid are subjected to equal forces of attraction on all sides, at the surface these forces act on one side of the molecules only, and therefore tend to pull them inwards. This causes the surface to pull itself together so as to occupy the least possible area, and it is this force which constitutes surface tension. The surface behaves as if stretched. There are various simple experiments that reveal the presence of surface tension. If a film is made on a loop of wire by dipping it in soap solution, a fine silk thread can be floated in the film, so that it forms a loop that is quite loose. If the portion of the film inside the loop is destroyed by touching it with filter paper, the film will break in the loop, which will now be pulled into a circular shape by the tension of the film around it (Fig. 17).

For the measurement of surface tension, various methods are used.

The size of drops of liquid falling from an orifice is dependent on surface tension; the larger the drops, the greater the surface tension. If the number of drops obtained by allowing a liquid to drop from a standard orifice in a given time is counted, we have a measure of the surface tension. Account must of course also be taken of the specific gravity of the liquid. The instrument used for this purpose is called a *stalagmometer* (Fig. 18). Another method depends on the fact that the height to which a fluid rises in a capillary tube is dependent on surface tension (and inversely on the diameter of the capillary). The difference in the heights to which two liquids rise in capillary tubes of known bore permits us to compare their surface tensions, and if this is known for one of the solutions, it can be determined for the other.

Besides existing between liquid and air, surface tension also exists at the interface between two immiscible liquids, and at that between sus-



Fib. 18.—Traube's stalagmometer. The surface tension is proportional to the number of drops formed in a given time. The right-angled tubes are for thin liquids, and the straight one for blood and other viscous fluids.

pended solid particles and liquid, as in colloidal solutions. Since, as we have seen, the surface area (interface) is enormously increased in these solutions, a very great surface energy is present, for this is equal to the surface tension multiplied by the surface area.

ADSORPTION

The surface tension between *liquid and air* is lowered when organic substances are dissolved in the liquid, but is slightly raised when inorganic salts are dissolved. The degree of lowering varies markedly according to the organic substance dissolved, being very pronounced with bile salts, upon which fact the well-known (Hay) test for the presence of bile in urine is based. Between *liquid and liquid,* as well as between

solid and liquid, the surface tension *is always lowered by dissolving substances in the liquid.* Now, at the interfaces between solid particles and liquid there must be a local accumulation of free surface energy, which will be equal to the surface tension multiplied by the surface (interface) area. A constant tendency exists for such free energy to be decreased and, since dissolved substances have this effect, they will become concentrated at the interface. This is known as *the principle of Willard Gibbs,* and it is of fundamental importance to the biochemist, because on it depends the phenomenon known as adsorption, which in the case of colloidal solutions may therefore be defined as the local concentration or condensation of dissolved substances at the interface between the two phases. The amount of substance concentrated at the interface can be calculated by a formula which takes into account the concentration of the dissolved substance, the temperature, and the surface tension at the interface (the Gibbs formula). After absorption has occurred, various reactions of a chemical, electrical or purely physical nature (e. g., diffusion) may follow at a rate which depends on the amount of the condensation.

Every-day Reactions Which Depend on Adsorption

1. Decolorization of liquids by charcoal. That no chemical reaction occurs in such a case is readily shown by the ease with which the pigment can be extracted from the charcoal.

2. Adsorption of gases by such solids as charcoal and spongy platinum. In these cases there must be great condensation, even a liquefaction of the gas, during which heat must be evolved. By adsorbing oxygen and hydrogen, spongy platinum causes these gases to combine and form water. The hemoglobin of blood may take up oxygen by a similar process.

3. Formation of solid surface films on solutions of protein, saponin, etc. The condensation may lead to coagulation, which explains why, if the froth produced by beating the white of an egg is allowed to stand, it can not be again beaten into a froth, the albumin having gone out of solution by surface coagulation.

An interesting phenomenon depending on the surface tension occurs when the protoplasmic contents of a ciliated infusorian is pressed out in water. A new membrane forms on the protoplasm because of surface concentration of all constituents which lower surface energy. By application of the principle of Willard Gibbs, A. B. Macallum[18] concludes that not only adsorption, as exhibited in a colloidal solution, but also the local accumulations of material often seen in cells, are associated with changes in surface energy. His conclusions are based largely on microscopic studies of various forms of cell exhibiting different degrees and types of activity,

and ingeniously stained for potassium by cobalt hexanitrite. By such a means the potassium stains intense black. In vegetable cells, local accumulations of potassium occur either near the interface between the clear and the chlorophyl-containing parts of the cell (spirogyra) or under a portion of the cell wall from which later a protrusion grows out to form the first stage in conjugation. The outgrowth from the cell, as well as the accumulation of potassium, may be the result of a low surface tension. In unicellular animal organisms, such as Vorticella, much less potassium is present, being confined to the base of the cilia, which Macallum believes indicates that the structures are produced as an outcome of low surface tension.

In the cells of higher animals, deposits of potassium are also localized; in striated muscle, for example, they occur in a zone at each end of the doubly refractive band and immediately adjacent to the singly refractive band. Changes in surface tension, associated with changes in the distribution of potassium, are believed by many to be responsible for muscular contraction. In nerves and nerve cells, potassium is concentrated at the axon and at the surfaces of the cells. Interesting suggestions are offered to explain the relationship among changes in surface tension at the terminations of axons (synapses, terminations in gland and muscle cells) brought about by the nerve impulse acting as a change in electric potential. Surface condensation of potassium has also been observed at the lumen border of gland cells (pancreas), and on the lumen surface of the cells of the renal tubules. Such observations indicate in what way surface tension may be called into play to control cellular activities. The field is new and almost unexplored, but there is already much to indicate that surface energy plays a most important role in the performance of many cellular activities.

Conditions That Influence or Are Influenced by Adsorption

Electric Changes.—Besides mere concentration, other forces come into play to assist or retard adsorption. One of the most important of these is electrical. Most solids when present as particles in a fluid carry a negative charge of electricity, some a positive one. In conformity with the Willard Gibbs law, a constant tendency will exist for this free energy to be diminished by the neutralization of the electric charge. This can occur by deposition on the interface of other particles carrying an electric charge of opposite sign or by the action of that present on ions. Charcoal in suspension in water, for instance, has a negative charge. If colloidal iron, which has a positive charge, is added to the solution, it will become deposited on the charcoal, as will also the cations of an inorganic salt. On account of electric adsorption, dyestuffs and bile

salts are adsorbed much more freely than they would be if the process depended solely on surface condensation; that is, if the Gibbs formula is used to calculate the adsorption, it will give values that are much below those actually found.

If the dissolved substance and the particles both have the same electric sign, adsorption will not occur. Filter paper, for example, has a negative charge and can not therefore adsorb a negative dye such as congo red (as shown by the depth to which it becomes stained); whereas it readily adsorbs night blue, which is positively charged. If the negative charge of the paper is lowered, it becomes capable of adsorbing some of the negative congo red. This can be effected either by placing the paper in alcohol or by adding inorganic salts (NaCl) to the water with which it is in contact. The positive-charged ions of Na, produced by dissociation, neutralize some of the negative charge on the paper, and allow a certain amount of adsorption of the negative-charged congo red to occur. As would be expected, acids and alkalies are capable of greatly altering the electric charges by the H and OH ions which they contribute.

Chemical Forces.—If the nature of the phase at the surface of which adsorption occurs is such that it can enter into chemical combination with the substance adsorbed, reactions will occur that do not obey the laws of mass action. By adsorption, reactions of a certain type may be encouraged over other reactions, even although the necessary reacting substances may be present in the solution (specific adsorption). The adsorbing substance itself is not, however, usually susceptible of chemical change even when it exists as very minute particles, as in the case of colloidal solutions. Nevertheless, adsorption may accelerate chemical reactions by bringing together in concentrated form substances of high chemical reactivity. In such cases the adsorbing substance itself does not enter into the chemical reaction, and can be recovered at the end in an unchanged condition. It acts as a catalyst (page 72). As we shall see later, enzymes act in this way—i. e., their rate of reaction is controlled by adsorption.*

The distinguishing feature of such adsorption phenomena is that a curve of the reaction (drawn by plotting amount of chemical change

*Another instance of the influence of surface energy on the course of chemical reactions is seen in the accelerative influence of charcoal on such reactions as the oxidation of formic acid, glycerol, etc. Surface tension may also cause retardation of chemical reactions, as is seen in the turbidity (due to the separation of chloroform) which gradually develops when a $\frac{M}{1}$ Na_2CO_3 solution is mixed with a $\frac{M}{2}$ chloral hydrate solution. The surface remains clear, because surface energy has prevented the reaction.

An important effect of surface tension on chemical reactions is also seen in the relationship between it and the absorption coefficient of gases (volume of gas dissolved by unit volume of liquid). The lower the surface tension, the greater the solubility of the gas. Oxygen and nitrogen are, for example, much more soluble in alcohol, hydrocarbons or oil than in water. This shows the futility of attempting to prevent the loss of gases from fluids such as blood by covering them with oils or hydrocarbons.

against concentration of reacting substances) is a parabola, indicating that the laws of mass action (page 23) are no longer followed. In order to be able to determine whether some particular process—as, for example, a fermentation process, or the absorption of oxygen by blood—is caused by adsorption, we must compare its curves, constructed according to the same principles, with the typical adsorption curve. A formula may be used in constructing the curves. In arriving at this formula, two facts have to be remembered: (1) As adsorption proceeds and less and less of the free energy on the adsorbing surface remains to be neutralized, the reaction slows off, until equilibrium is reached. The more dilute the solution, the greater is the proportion of its contents to be adsorbed, which means that if *a* is the amount of substance adsorbed from a certain solution, then, from a solution of twice that strength, somewhat less than 2 *a* will be adsorbed—i. e., *a* multiplied by some root of 2. Although the formula is one belonging to the class known as parabolic, it must not be assumed that every reaction which happens to give such a parabolic curve (such as the combination of O_2 with hemoglobin under certain conditions) (see page 383) must be one dependent on adsorption.

It must be understood that although the substance that is removed from a solution by adsorption is no longer capable of contributing to the conductivity or the osmotic pressure of the solution, it is nevertheless not so firmly fixed that it can not be set free again by purely mechanical means, as by constant dilution of the fluid. If charcoal which has adsorbed sugar is placed in a dialyzer made of membrane the pores of which allow sugar but not charcoal to pass through, the sugar will gradually be removed if the dialyzer is immersed in running water. A certain equilibrium exists between the substance adsorbed and the same substance still remaining in solution. If the latter is constantly diminishing by dialysis, the adsorption compound must break down to maintain the equilibrium. It is clear, however, that the process of removal will be extremely slow. The ability of adsorbed substances to withstand removal by washing is taken advantage of by nature in holding back foodstuffs in the soil.

Physiologic Processes Depending on Adsorption

Instances in which adsorption undoubtedly plays a most important part in physiologic processes are as follows:

1. The action of enzymes (see page 71).
2. The combination of toxin with antitoxin occurs according to the laws of adsorption rather than those of mass action. In this case it is important to note that when the toxin of diphtheria is added in small suc-

cessive quantities to diphtheria antitoxin, more toxin is neutralized than when the toxin is all added at once. A similar phenomenon can also be observed by adding filter paper to congo red, more of the pigment being adsorbed when the paper is added in small quantities than when added all at once. The explanation is that relatively more adsorption of a given substance occurs from a dilute than from a strong solution (cf. page 69).

3. The sensitizing of leucocytes by opsonins, as well as the subsequent ingestion of bacilli by the sensitized leucocytes, both of which follow the course of an adsorption reaction.

4. The formation of adsorption compounds, such as the inorganic salts and proteins and the complex lecithin compounds that can be extracted from egg yolk or brain tissue. In such compounds the laws of chemical proportion no longer hold, and properties may be exhibited that are quite different from those of either one of its components. When yolk of egg is extracted with ether, for example, a compound of lecithin with vitellin goes into solution, although vitellin itself is quite insoluble in ether.* There can be no doubt that adsorption compounds of this character are very abundant in living cells, and that they are constantly being formed and broken down.

*By mixing solutions of egg albumin, congo red and a dye called fustic in the presence of alum, the colloidal particles of which each is composed run together to form larger colloidal aggregates, which by ultramicroscopic examination can be seen to be composed of a red, a yellow and a green colloidal particle. The attractive force holding the particles together is electric in this case.

CHAPTER IX

FERMENTS, OR ENZYMES

One of the most striking developments of modern research in biochemistry concerns the nature of enzyme action. So remarkable are many of the facts that have been brought to light that it can not fail to interest every one engaged in the study of life phenomena—whatever the nature of that study may be—to know something of the main questions at present occupying the attention of investigators in this field. In this chapter a brief survey will be given of some of these questions; no attempt will be made at completeness, and only where necessary for the sake of example will reference be made to individual types of enzyme action.

The discovery by Buchner that an enzyme can be expressed from yeast cells which is capable of instantly bringing about the alcoholic fermentation of dextrose solutions has been responsible for a great deal of the modern advance. Formerly, yeast cells were believed to bring about alcoholic fermentation as a result of their growth: it was believed to be a life phenomenon, or "vital process." Now we know that yeast cells produce an intracellular ferment or endo-enzyme* to which its sucroclastic properties are due and which can act apart from the cells that produce it. It is no great stretch of imagination to think of all chemical reactions mediated by cellular activity as due to a similar mechanism, and this thought has led to the hypothesis that all processes of intermediary metabolism in the animal and plant are caused by enzyme action. Before Buchner's day we knew only of the extracellular enzymes (such, for example, as the digestive ferments), that is to say, of enzymes, produced indeed by cells, but secreted from them and acting outside their protoplasm; now we must recognize intracellular enzymes acting where they are produced, in the protoplasm of the cell. But we must not permit this conception to carry us too far. Without further investigation we must not imagine that the riddle of life is thus solved.

As an example of the rôle which extra- and intracellular enzymes are supposed to play in the animal economy may be cited the metabolism of protein. Proteolytic enzymes are very widely distributed in the active tissues of the animal and plant. By their agency in animal life, the com-

*The terms "ferment" and "enzyme" are synonymous, but the latter is preferable as the noun, leaving the former to be used as the verb.

plex protein molecule is split up to render it absorbable from the intestine, and the tissues appropriate from the blood those of the degradation products that they require for the construction of protoplasm, which, later, they decompose so as to utilize the energy which the organism demands. All these processes are believed to be the work of enzymes.

The Nature of Enzyme Action

The changes brought about by enzymes can also be accomplished by ordinary chemical means, but these have often to be of a very energetic nature to accomplish what the enzyme can so quickly and quietly perform.

It is the custom to regard enzymes as *catalysts.* A catalyst is a substance which accelerates (or retards) a chemical reaction which in its absence could proceed at a different (usually slower) pace. The action of catalysts has been aptly likened to that of a lubricant. A weight placed at the top of an inclined plane, so held that the weight only slowly slips down, has its velocity greatly increased if its under surface be oiled. The oil accelerates the action but does not affect the ultimate result. Catalysts do not combine with the final products of the reaction, these being, as a rule, the same as they would have been had no catalyst been added. Another characteristic is the tremendous amount of chemical change which even a trace of catalyst can induce. There are many examples of catalysts in the inorganic world, among which may be cited the action of spongy platinum on hydrogen peroxide. This substance normally tends to decompose into water and oxygen, but if a small amount of spongy platinum is added to it, the decomposition is greatly accelerated: $H_2O_2 = H_2O + O$.

A very good example of the action of an inorganic catalyst is that of the hydrogen ion on cane sugar, or other disaccharides, in the presence of water. It accelerates the hydrolysis. Cane sugar solution at room temperature does not indeed, in sterile solution, undergo any appreciable hydrolysis, but at 100° C. it does, which leads us to believe that, though inappreciable, the action also occurs at room temperature. By adding a little hydrochloric acid, or other acid not having an oxidizing effect on sugar, we greatly accelerate the hydrolysis because of the hydrogen ions present in the acid solution. Within certain limits the rate of hydrolysis is proportional to the amount of catalyst present.

Enzymes, like other catalysts, produce their action when present in very small amounts (e. g., sucrase can hydrolyze 200,000 times its weight of cane sugar; diastase can convert starch to sugar in a dilution of 1-1,000,000) and there is a distinct relationship between the amount of enzyme present and the rate of the reaction. The final product of the

reaction is, however, the same at whatever rate it proceeds, and the enzyme does not appear in the final products. Many enzymes such as diastase can be found unaltered in amount after they have completed their action. This is determined by adding a fresh supply of substrate (that is, of material to be acted on), when the enzymic action proceeds again in the usual way. The same is no doubt true for all enzymes, though as yet it can actually be proved for only a few of them. Enzymes, therefore, may be defined as catalysts produced by living organisms.

The Properties of Enzymes

Although enzymes are examples of catalysts, they exhibit many properties that appear to differ from those of inorganic catalysts. It will, therefore, be advisable in considering each quality to compare it in catalysts and enzymes, for by this method a much clearer conception of the nature of enzyme action can be gained (Bayliss[19]). Those properties that are strictly peculiar to enzymes we shall consider later.

1. *Most enzymes are remarkably specific in their action, whereas inorganic catalysts are very much less so.* Thus, in the case of the enzymes which bring about inversion of disaccharides, this specificity is clearly shown. There is a special enzyme for each of the three disaccharides—maltose, lactose and cane sugar—and one of these can not replace another.

Still more strikingly is this specificity of enzyme action demonstrated in the fact that certain enzymes, such as zymase (expressed from yeast), will act only on bodies having a certain configuration, that is, having their side chains arranged in a certain way. Thus, there are two varieties of dextrose (α and β), which differ from each other solely in the fact that the side chains are arranged in different positions with relation to the central chain of carbon atoms. This form of isomerism is called stereoisomerism because the two bodies rotate the plane of polarized light to an equal degree in opposite directions. Zymase acts on one of these but not on the other, and there are innumerable examples of the same kind. Indeed, of all bodies that exist in two stereoisomers only one is found in living cells and it is on this variety alone that the enzymes in animals can act. A similar specificity exists between certain drugs and their pharmacologic action.

Specificity of action is explained by supposing that a union occurs between the substrate and the enzyme, and for this union to take place the enzyme must possess a configuration which corresponds accurately with that of the substrate. The process has been compared to a lock and key; the key must be shaped to fit the lock, or it can not operate. The specificity does not, however, in itself disprove the close

relationship between enzymes and inorganic catalysts, for on the one hand there are several enzymes which do not exhibit this property, and on the other, there are inorganic catalysts which do. For example, lipase, the fat-splitting enzyme of pancreatic juice, decomposes not only fats but to a greater or less degree a number of bodies of the same general build (esters), and tyrosinase can decompose, not tyrosin alone, but all phenol compounds. Conversely, the hydrogen ion—to the presence of which acids owe their catalytic powers—can decompose the ordinary esters (that is, of acids containing the carboxyl or COOH group) but it has no action on the sulphonic esters. However, enzymes are certainly much more specific in their action than inorganic catalysts.

2. *Temperature does not influence catalysis and enzyme action in the same way.* As the temperature is raised in the case of inorganic catalysts, the reaction becomes about doubled in rapidity for each rise of 10° C., whereas in the case of enzymes it becomes increased up to a certain *optimum temperature*, beyond which, as the temperature rises, the reaction is first slowed and then disappears altogether.

This peculiarity of enzymes as compared with inorganic catalysts need not in itself disprove the analogy between the two, because enzymes do not form true, but colloidal solutions. Colloidal solutions, as we have seen, are really fine suspensions of ultramicroscopic particles; there is no splitting into ions of the dissolved substance, as is the case with true (molecular) solutions, but the colloid is suspended in the water or other solvent to form a heterogeneous system (page 51), on which account the surface area of the menstruum is enormously increased. Rise in temperature alters the extent of the surface area, and thereby introduces an influence which progressively opposes catalysis.

Although inorganic catalysts in molecular solution show no optimum temperature but increase in activity in proportion as the temperature is raised, inorganic colloidal catalysts may show an optimum temperature. Thus, spongy platinum shows an optimum temperature in its action on a mixture of hydrogen and oxygen. It has therefore been suggested that it is because they are colloids that enzymes exhibit this property.

3. *Inorganic catalysts frequently carry the reaction to a further stage than that attained by the action of enzymes.* For example, acid breaks down the protein molecule much more completely than do the proteolytic enzymes. This difference is perhaps explained by the fact that enzymes are retarded in their activities when there comes to be a certain accumulation of the products of the reaction present. The final stages in the reaction may become so slow as to be almost inappreciable. This decrease in activity is partly due to a union between the enzyme and the products of its activity.

4. *The velocity constant in the case of inorganic catalysts remains unchanged throughout the reaction, whereas in the case of enzymes it becomes either less or greater as the process proceeds.* When a substance is changed by catalytic action, it is, of course, constantly being diminished in concentration so that less and less of it remains to be acted on. This implies that there are fewer molecules present for the same amount of catalyst to act on and consequently that the amount changed in a unit of time is progressively less. At any moment, therefore, the rate of catalysis will be proportional to the amount of substance (substrate) left. To understand this we must refer back to what we have learned about mass action. If we suppose that two substances A and B react to form two other substances C and D, then, by the law of mass action, the reaction will not go on to completion but will stop when a certain equilibrium is reached. The reaction can be represented by the equation $A + B \rightleftarrows C + D$, which means that it proceeds at a rate proportional to the reacting molecules. In some cases this reaction goes on until either A or B has practically disappeared (that is, the equilibrium point is very near the right of the equation), as is the case in the inversion of cane sugar:

$$C_{12} H_{22} O_{11} + H_2O = C_6 H_{12} O_6 + C_6 H_{12} O_6$$

Taking place as it does in an excess of water, and there being very little tendency for this reaction to go in the opposite direction (cf. reversible action) (page 25), the only thing which will influence its velocity is the concentration of cane sugar; in other words, the velocity of the reaction at any moment will depend solely on the concentration, C, of the material still left undecomposed. This can be determined by means of an equation.*

The value of such an equation is that it gives us a figure K, representing the amount of inversion that would occur in each unit of time if the cane sugar were kept in constant concentration. When, for example, it is stated that K for a particular strength of acid acting on cane sugar solution is 0.002, this means that when volume, concentration of acid and

*If x be the amount of sugar inverted in time *t*, and if we use a figure called a constant (*K*) to express the fundamental rate of the reaction (which will therefore be different for different reactions), then $\frac{x}{t} = KC$. But C can not be the same at any two consecutive periods of time, because the reaction is going on continuously. This renders it necessary to use the notation of the differential calculus, and we have $\frac{\delta x}{\delta t} = KC$. The sign δ indicates that the reaction is a constantly changing one so that δx and δt represent such infinitely small amounts that they can not be measured. By methods of integration, however, it can be shown that the above equation may be written:

$$K = \frac{1}{T_2 - T} \text{ log. nat. } \frac{C_1}{C_2},$$

thus permitting us to find the value of K (C_1 C_2 being the concentrations of cane sugar at the times T_1 T_2).

Any two determinations during the course of the reaction can be used for calculating K. These equations apply only to cases in which but one substance is changing (monomolecular reaction). When two substances are involved, the equation is more complicated.

temperature are constant in a gram-molecular solution of sugar, 0.002 gram-molecule of sugar would be inverted the first minute and 0.002 gram each succeeding minute, provided we could keep the solution constantly a gram-molecular one, that is, provided we could add sugar just as quickly as it becomes inverted.

At first sight it may appear of little practical importance to determine K. In our present discussion concerning the nature of enzyme action, it is however of great value for, whereas with inorganic catalysis K is really of constant value, with enzyme action it is not so. Thus, when cane sugar is inverted by sucrase—an enzyme present in the intestine and in yeast—the constant gradually rises; for most other unimolecular reactions mediated by enzymes it gradually falls; for example, the action of trypsin on proteins.

Where there is a great excess of substance to be acted on, in comparison with the amount of enzyme present, it will be found that a more constant value than K is obtained when we compute the absolute amount of substance decomposed in a given time. In such a case, too, the amount of change in a given time will be proportional to the amount of enzyme present, indicating that some sort of combination between enzyme and substrate must be the first step in the fermentative process. This fact has been noticed by us in connection with the hydrolysis of glycogen in the liver. When there is an excess of glycogen present, the amounts which disappear in equal intervals of time after death are the same; when, on the contrary, there is not much glycogen, the amount which disappears gradually declines, but, if K be computed by the above equation, it is constant.

To make these facts clear it may be well to pause for a moment to consider an illustration. The conditions obtaining when there is a large excess of substrate over enzyme may be compared to those governing the removal of a pile of bricks from one place to another by a number of men. The pile of bricks represents the substrate; the men, the enzyme. If each man works up to his capacity, it is plain that the number of bricks transferred in a given time will not depend at all on the size of the pile to be transferred. When, however, the pile of bricks gets small, though the same number of men continue to work the number of bricks transferred in a given time falls off, because the men interfere with one another's activities in securing their loads from the pile. When a similar stage is arrived at in enzyme processes, we have to use the velocity constant to show how much work could be done by the enzyme if the amount of substrate were maintained of constant amount.

In the large volume of recent work which has been done with the object of discovering the cause of these variations in the velocity con-

stant in the case of enzymes, four important conditions have been recognized: (1) reversibility; (2) gradual destruction of the enzyme; (3) combination of the enzyme with products of the reaction; (4) autocatalysis.

Of these four influences the only one which could be held accountable for an increase in the activity of the enzyme is *autocatalysis;* in this process the enzyme by its action produces substances which intensify its own activity. In some cases at least—for example, the action of invertase on cane sugar—these are acid bodies, a moderate increase in acidity favoring the action of this enzyme.

The other influences all tend to retard the reaction and progressively lower the value of K. Negative autocatalysis occurs when the enzyme produces products which interfere with its activity. Gradual destruction of the enzyme and its union with the products of its activity will manifestly also decrease its power. There is plenty of evidence that both of these processes may occur.

Reversibility of Enzyme Action

But the most important of all the causes of retardation of enzyme activity is undoubtedly *reversibility of action,* which is an application of the law of mass action (page 25). If we take the saponification of an ester, the equation is:

$$\underset{\text{(ethyl butyrate)}}{CH_3CH_2CH_2COOC_2H_5} + H_2O \leftrightarrows \underset{\text{(butyric acid)}}{CH_3CH_2CH_2COOH} + \underset{\text{(ethyl alcohol)}}{C_2H_5OH}.$$

The equilibrium point is not so near the position of complete hydrolysis as in the case of the inversion of saccharose; in other words, the tendency for the bodies produced by the hydrolysis to reunite and form the original substances is quite marked, so that the reaction comes to an end before all the ethyl butyrate has been decomposed. For some time before the equilibrium point is reached, there will have existed a progressively increasing opposition to the breakdown of the ester, as a consequence of which, when enzymes are used to accelerate the reaction, the velocity constant, as determined by the above equation, will gradually fall as the reaction proceeds. Conversely, in a mixture of ethyl alcohol and butyric acid there is very slow synthesis to ethyl butyrate, and here again lipase accelerates the process; it induces a recognizable synthesis within a short time. Ethyl butyrate is usually employed for these experiments because, on account of its odor, the ester is readily recognized. Thus, if the alcohol and acid be mixed alone, no ester will be detectable, but if some lipase be added, it will soon become so. Similar synthetic action of lipase has also been demonstrated for mono- and tri-olein.

It should be clearly understood that pure catalysts, such as the hydrogen ion, in accelerating a reaction like the above, do so equally in both directions, so that the position of equilibrium remains unchanged. Enzymes may, however, cause this position to change because of their forming intermediate combinations.

The reverse phase of certain reactions is probably the cause of at least some of the synthetic processes which occur in the animal body. A great difficulty in accepting such a view, however, is the fact that the equilibrium point of all hydrolytic reactions, in the presence of an excess of water, is so near complete hydrolysis that very little synthesis can be possible. That is true so long as the substance synthesized is soluble, but if it is nearly insoluble in water, or if it is immediately removed from the site of the reaction by diffusion, or in any other way, then it is obvious that it will go on being synthesized by the reaction. Thus, in the intestine neutral fat is hydrolyzed by pancreatic lipase into fatty acid and glycerin, which are absorbed into the epithelium, where they again come under the influence of intracellular lipase. This latter will tend to accelerate the synthesis of neutral fat from the fatty acid and glycerin until the equilibrium point of the system (fat acid + glycerin $\rightleftarrows$ neutral fat + H_2O) is again reached; but this point, although it is near the right hand of the equation, will really never be reached for the reason that the neutral fat, as quickly as it is formed, will become deposited in insoluble globules in the protoplasm and thus be removed from the equation. In support of this view it has been found that lipase is present in intestinal mucosa after all traces of adherent pancreatic juice have been washed away. By similar reactions the fat of the tissues becomes decomposed to fatty acid and glycerin and passes out of the blood when the concentration of fat in this fluid falls below a certain level. Provided one of the substances synthesized is insoluble or can in some other way be removed from the reaction, it is plain that, even though the equilibrium point is very near to that of complete hydrolysis, yet the reversion will be sufficient to do all that is required of it.

Results such as the above have prompted many to conclude that it is by such reversible action that all synthetic processes occur in the living organism. But the demonstrable synthesis of an ester must not be taken as evidence that all other syntheses are explainable on the same basis. For example, we have seen above that in the case of cane sugar the equilibrium point in the equation is so near that of complete hydrolysis, that no measurable amount of cane sugar is formed when dextrose and levulose are allowed to act on each other, and that cane sugar does not appear when sucrase is added to the mixture. If instead of sucrase we take another of the sugar enzymes—namely, maltase, which accelerates the

decomposition of maltose into two molecules of glucose—there is, however, evidence of synthesis as a result of the acceleration of a reversible reaction. To understand these results we must remember that ordinary dextrose is a mixture of two stereoisomers designated α and β. When two molecules of α dextrose condense (that is, fuse togther with the loss of a molecule of water) maltose is formed, but when two molecules of β dextrose condense isomaltose results. There is some controversy as to whether *maltose* is really responsible for the synthesis of α dextrose molecules to maltose, it being claimed by some that this is accomplished by another enzyme, *emulsine.* If this were true it would materially minimize the importance of reversible action as a factor in cellular synthesis. The latest evidence goes to show, however, that it is maltase and not emulsine that is responsible in the above case (cf. Bayliss).

Evidence, both direct and indirect, is also steadily accumulating to show that enzymes may accelerate the synthesis of proteins. As pieces of direct evidence we have: (1) the retardation of the digestive action of trypsin, etc., which sets in after the process has gone on for a time, and (2) the recommencement of a digestive process apparently at an end, if the products of the digestion are removed by dialysis or other means. As direct evidence may be cited the formation of synthetic products when pepsin is added to concentrated solutions of peptone, and the diminution in the number of small molecules, as judged by measurements of electrical conductivity, when trypsin is added to the products of tryptic digestion of caseinogen. Protamine—a simple form of protein—has also been found to be produced when trypsin—obtained from a mollusc—was added to a tryptic digest of the same protamine. The significance of these facts in connection with the metabolism of the amino aids will be evident when we come to study this subject (page 598).*

Specificity of Enzyme Action

Although in all of the above features of enzyme action there is nothing to contradict the view that they are catalytic agents, there remains one peculiarity which at first sight seems uninterpretable on such a basis. This is with regard to their often remarkable specificity of action. Thus, as we have seen, maltase can hydrolyze maltose alone (which is composed of two α-dextrose molecules), but not isomaltose (composed of β-dextrose). This means that mere difference in the configuration of molecules is sufficient to alter the influence of enzymes on them. Since such differences could not influence that of inorganic catalysts we must

*We have been unable in this laboratory to demonstrate any synthesis of glycogen when glycogenase is added to a hydrolysis mixture of dextrine, maltose and glucose produced by the prolonged action of glycogenase on pure glycogen.

explain the cause of the difference. This has been done on the basis either that enzymes are colloids or that the active (catalytic) group of the enzyme is attached to a colloid molecule. Before a substance can be acted on, it must combine with the colloid, which it does by the process of adsorption (see page 65). This can occur, however, only when there is a harmony between the adsorbing substance and the substance adsorbed. Instances of the specificity of adsorption have already been given.

In support of this view it has been found that of the two *proteases*, α and β, in the spleen, one is adsorbed but not the other when a solution containing them is shaken with Kieselguhr. Furthermore, when solutions of invertase are shaken with certain inert powders, the invertase is adsorbed by some of them but not by others. In strong support of the adsorption hypothesis is also the fact that the same mathematical laws as apply in the process of adsorption are obeyed in the ratio which exists between the activity of an enzyme and its concentration in the solution.

To sum up, then, catalysis as exhibited by enzymes involves three processes: (1) contact between the enzyme and the substrate, which will be dependent on their rates of diffusion; (2) adsorption between them, which will depend on their configurations (cf. the lock and key simile); and (3) the chemical change which itself probably takes place in two stages. In connection with the third process, it is probable that an initial compound of a definite chemical nature is first formed, followed by the hydrolytic or other chemical change, after which the enzyme group becomes free.

It is very significant in this connection to note that in their solubilities there exists a distinct relationship between the ferments and the substrates on which they react. Thus, trypsin is very soluble in water and acts on water-soluble proteins; lipase is soluble in fat solvents.

Certain Peculiarities of Enzymes

Notwithstanding the very strong case that is made out for the catalytic hypothesis, there are certain facts which many find it difficult to make conform with such a view. One of these is that dextrose can undergo three distinct and separate types of decomposition according to the enzyme allowed to act on it. These are alcoholic fermentation, butyric acid fermentation and lactic acid fermentation. It is difficult to see how simple catalytic action can be responsible for all three results. The enzyme must not only initiate the changes but also direct their course.

Another peculiarity is that when certain enzymes—e. g., rennin, pep-

sin, etc.—are inoculated in animals, they cause specific *antienzymes* to appear in the blood of the inoculated animal. Thus, when antirennin serum is added to milk it greatly hinders clotting on the subsequent addition of rennin. It is probable that powerful antienzymes are produced in the animal body for the purpose of protecting the tissues from attack by enzymes. It is on account of the presence of antienzymes that intestinal parasites can exist in the intestine, and the immunity from digestion which the mucosa of the gastrointestinal tract enjoys, is believed to be due to the same cause. But there is considerable doubt regarding this claim. Fresh pancreatic juice when injected into the empty intestine digests its walls. When food is present in the intestine it evidently prevents digestion of the walls by diverting the enzyme to itself.

Types of Enzyme

Having learned something about the general nature of enzyme action, we may now turn our attention to certain details that have a practical importance. In the first place, with regard to nomenclature, in the earlier work each newly discovered enzyme received a name which was often quite inappropriate. Many of these names are retained, such as pepsin, trypsin, ptyalin, etc., but it is now customary to name the enzyme according to the substance on which it acts. This is done either by replacing the last part of the name of the substance acted on by the termination -ase (for example, the enzyme which inverts maltose is called maltase), or by merely adding -ase to the name of the substance acted upon (thus, the enzyme which hydrolyzes glycogen is called glycogenase).

Most of the enzymes in the animal body accelerate hydrolytic processes and are classified according to the chemical nature of the substrate on which they work. Thus, we have:

1. The *amylases*—accelerating the hydrolysis of polysaccharides, e. g., ptyalin (in saliva), amylopsin (in pancreatic juice), glycogenase (in liver), diastase (in malt).

2. The *invertases*—accelerating hydrolysis of disaccharides, e. g., maltase, lactase and sucrase (in succus entericus).

3. The *proteinases*—accelerating hydrolysis of proteins, e. g., pepsin (in gastric juice), trypsin (in pancreatic juice), erepsin, intracellular proteinases.

4. The *lipases*—accelerating disruption of neutral fats, e. g., steapsin (in pancreatic juice), intracellular lipases.

5. *Arginase*—accelerating hydrolysis of arginin into urea and ornithin, (intracellular).

6. *Urease*—accelerating hydrolysis of urea to ammonium carbonate (in many microorganisms and in the soy bean).

7. *Glyoxylase*—converting glyoxals into lactic acid (page 666).

Other enzymes accelerate oxidative processes and are called *oxidases* and *peroxidases*. Others bring about the displacement of an amino group by hydroxyl (*desamidases*). Others cause coagulation (*coagulative ferments*), e. g., thrombin, rennin. One of the enzymes present in succus entericus acts by converting the zymogen (trypsinogen) into the enzyme (trypsin).

Enzyme Preparations

So far it has been impossible to prepare enzymes in a pure state although, being colloidal in nature, they are readily precipitated or adsorbed along with other colloids.

Since most enzymes exist in cells, it is necessary to break up the cells in order to isolate the enzyme. This is done in various ways. By one method the cells are ground in a mortar with fine sand, then made into a paste with infusorial earth (Kieselguhr), the paste enclosed in stout canvas and placed under an hydraulic press at about 300 atmospheres pressure; a clear fluid separates and this contains the enzymes. Another way is to freeze the tissue with liquid air and grind it in a steel mortar by means of a machine. Still another and less expensive method, and one which we have found most useful for organs and tissues, consists in reducing the tissue to a pulp and, after sieving it to get rid of connective tissue, etc., spreading the pulp on glass plates and drying in a slightly warmed, dry air current. The scales of dried material are then ground in a paint mill with toluene, and the resulting suspension filtered; the powder which remains on the filter, after thorough washing with toluene, is dried and kept for future use. The toluene removes all the fatty substances, so that when shaken with water, etc., the enzymes dissolve.

Conditions for Enzymic Activity

Reactions brought about by intracellular enzymes are very readily inhibited when there comes to be a certain accumulation of their products of action. Thus, yeast ceases to ferment sugar when the alcohol has accumulated to a certain percentage. This action is partially due to a toxic action of the alcohol on the cell, which paralyzes its power of absorbing the substance to be acted on by the intracellular enzyme. If these products be not in some way removed, they will ultimately kill the cell and stop the fermentation. We have seen above how the accumulation of products may interfere with the activities of enzymes in

other ways in which the enzyme does not suffer destruction, as is shown by the fact that it resumes its original activities on removal of the products.

Enzymes, both intracellular and extracellular, are very sensitive towards the inorganic composition of the medium in which they are acting. For the intracellular enzymes this is what we should expect when we bear in mind the profound influence of inorganic salts on the heart beat and on cell growth and division. This influence of salts and of reaction (acidity, etc.) on the life of the cell is so pronounced as to lead some observers to believe that abnormal cell multiplication in the body, as in the case of tumor formation, is due to changes in the inorganic composition of the tissue fluids. Extracellular enzymes are also very susceptible to the influence of inorganic salts but more especially so towards the reaction of the solution. In terms of modern chemistry we may say that the concentration of $H^{\cdot}$ and OH' ions has a profound influence on the activities of enzymes. Most of the enzymes of the animal body perform their action normally in the presence of a slight excess of OH' ions, that is, in faintly alkaline reaction. Indeed the only exception of importance to this is the pepsin of gastric juice, which normally acts in an acid medium. An excess of either OH' or $H^{\cdot}$ ions inhibits the activity of the enzyme and usually destroys it permanently. The activities of enzymes are also influenced by light, many of them being destroyed by sunlight; cells such as microorganisms are similarly affected.

Before being secreted the digestive enzymes exist in the cells which produce them as inactive precursors called *zymogens*. The granules seen in resting gland cells are of this nature. The activation of the zymogen, or its conversion into the enzyme, occurs after it has left the cell, and this has been considered as another safeguard to digestion of the cell. Sometimes the activation does not occur until the zymogen has travelled some distance along the gland duct, as in the case of the proteolytic enzyme of pancreatic juice. Till it reaches the intestine, this exists as trypsinogen (the zymogen), but it is here acted on by another enzyme-like body produced by the intestinal epithelium and called enterokinase.

PHYSICOCHEMICAL REFERENCES

(Monographs and Original Papers)

[1]Bayliss, W. M.: Principles of General Physiology, Longmans, Green & Co., 1915.
[2]Philip, J. C.: Physical Chemistry, Its Bearing on Biology and Medicine, Arnold, ed. 2, 1914.
[3]McClendon, J. S.: Physical Chemistry of Vital Phenomena, Princeton University Press, 1917.
[4]Starling, E. H.: Principles of Human Physiology, ed. 2, 1915, Lea and Febiger.

[5]Kahlenberg, L.: Jour. Physical Chem., 1906, x, 141.
[6]Reid, E., Weymouth: Jour. Physiol., 1898, xxii, lvi.
[7]Wilson, T. M.: Am. Jour. Physiol, 1905, xiii, 150.
[8]Haldane, J. S., and Priestley, J. G.: Jour. Physiol., 1916, l, 296; Priestley, J. G: Ibid., p. 304.
[9]Clark, W. M., and Lubs, H. A.: Jour. Bacteriology, 1917, ii, 1 and 109.
[10]Henderson, L. J.: The Excretion of Acid in Health and Disease, Harvey Lectures, J. B. Lippincott Co., 1915, x, 132.
[11]Henderson, L. J.: The Fitness of the Environment, Macmillan, N. Y., 1913.
[12]Van Slyke, D. D.: Jour. Biol. Chem., 1917, xxx, 289, 347.
[13]Levy, R. L., and Rowntree, L. G.: Arch. Int. Med., 1916, xvii, 525.
[14]Cullen, G. E.: Jour. Biol. Chem., 1917, xxx, 369.
[15]Palmer, W. W., and Henderson, L. J.: Arch. Int. Med., 1913, xii, 153.
[16]Sellards, A. W.: The Principles of Acidosis and Clinical Methods for Its Study, Harvard University Press, Cambridge, 1917.
[17]Lloyd, F. H.: Private communication.
[18]Macallum, A. B.: Surface Tension and Vital Phenomena. University of Toronto Studies, No. 8, 1912; also Ergebnisse der Physiologie, 1911, ii, 598.
[19]Bayliss, W. M.: Enzymic Action, ed. 2. Monographs in Biochemistry, Longmans, Green & Co.

PART II

THE BLOOD AND THE LYMPH

CHAPTER X

BLOOD: ITS GENERAL PROPERTIES

BY R. G. PEARCE, B.A., M.D.

The blood, being the carrier of the nutritive and waste substances of the body's metabolism, must at one time or another contain all the materials which compose the tissues in addition to those which are peculiar to the blood itself. It is a very complex fluid, and all of its constituents are not fully known. Structurally it is composed of water in which are dissolved various gases and organic and inorganic bodies, the corpuscles and platelets.

THE QUANTITY OF BLOOD IN THE BODY

The most accurate method of determining the volume of blood in the body is by bleeding and subsequently washing out the blood from the vessels and then estimating the amount of hemoglobin in the total fluid (Welcher's method). This method employed in the case of two criminals who had been decapitated gave the weight of the blood as 7.7 and 7.2 per cent of the body weight. Bloodless methods for determining the total volume of blood are based upon the principle of adding a definite quantity of a known substance to the circulation and then estimating its concentration in a sample of blood withdrawn from the body shortly afterward. If the substance can not leave the blood vessels and does not cause fluid to be withdrawn from the tissues, the total quantity of blood in the body can be calculated from the concentration of the injected substance in the blood. The most accurate methods based on this principle are Haldane and Smith's, in which carbon monoxide gas is inhaled in a given amount and the carbon monoxide hemoglobin subsequently determined colorimetrically; and Keith, Rowntree and Geraghty's, which employs vital red, a dye of low diffusibility. The dye remains long enough in the body to be thoroughly mixed with the blood, and its concentration in the plasma is determined colorimetrically

by comparing with a suitable standard mixture of dye and serum. These methods give the total amount of blood in the body as from 5 to 8.8 per cent of its weight. Meek has recently developed a method in which gum acacia is used. After mixing with the blood, the concentration of this substance is determined from the calcium content. Being colloid, none of the gum leaves the blood vessels.

The newer methods have shown that the volume of the circulating fluid is maintained fairly constant in spite of influences tending to alter it. The body accomplishes this by drawing upon the reserve fluid in the tissues and by varying the rate of water excretion, particularly through the kidneys. Years ago the doctrine of an increased amount of blood in the body (plethora) gave rise to the therapeutic use of bleeding. Especially was this thought to be useful in conditions which we now recognize as chronic hypertension, and which show no increase in blood volume. Indeed variation in blood volume is not common, although plethora may occur in polycythemia, chlorosis, and anemias, and there may be a temporary reduction in the amount of blood in diseases in which there is a great depletion of water, as in Asiatic cholera, and following very severe hemorrhage.

While the total quantity of the blood in the body does not vary greatly, the concentration of its various constituents is subject to distinct change. The volume percentages of the corpuscles and the plasma can be approximately determined by allowing oxalated blood to sediment or by centrifuging in a graduated cylinder by the use of the hematocrit. Such methods are not very reliable, but may yield some important information. Normally 45 to 50 per cent of the volume of blood is composed of corpuscles. It varies more or less directly with the number of red blood cells.

THE WATER CONTENT OF THE BLOOD

Since the blood plasma is essentially a watery solution, some idea of its water content can be obtained by a determination of the specific gravity. The most accurate method for accomplishing this is to determine directly the weight of a given volume of blood and compare it with the weight of the same volume of water. Since this method requires a rather large amount of blood, indirect methods using smaller amounts have been devised. One of these (Hammerschlag's) uses a solution of chloroform and benzol of a specific gravity of about 1.050, in which a drop of blood is suspended by delivering it cautiously from a pipette bent at right angles near its tip. If the drop sinks, chloroform is added; if it rises, benzol is added until the drop remains suspended.

The specific gravity of the benzol-chloroform mixture is then determined, and this value is supposed to give the specific gravity of the blood.

The specific gravity of the blood determined in this way varies between 1.040 and 1.065. It is somewhat less after eating and increases after exercise; it is slightly lower during the day than at night, and the variation in individuals is considerable. The changes which occur in the specific gravity of the blood in disease are chiefly due to variation in the percentage of protein, since the salt content of the blood is relatively fixed. It is only when great changes occur in the concentration of the noncolloidal salts that they markedly affect the specific gravity.

From 90 to 92 per cent of the plasma and from 59.2 to 68.7 per cent of the corpuscles consist of water. Of the whole blood, from 60 to 70 per cent by volume or about 55 per cent by weight consists of plasma; and from 40 to 30 per cent by volume or 45 per cent by weight consists of corpuscles.

THE PROTEINS OF THE BLOOD

The plasma obtained by centrifuging the blood rendered noncoagulable by oxalates, hirudin or other means (see page 99), contains 5 to 8 per cent of coagulable proteins. These proteins are serum albumin, serum globulin, and fibrinogen. They can be separated from each other by the use of acids and neutral salts. Their proportion varies under different conditions, but is approximately as follows:

Fibrinogen	0.15-0.6%
Serum globulin	3.8%
Serum albumin	2.5%

The amount of fibrinogen is subject to the greatest variation (Mathews).

Fibrinogen

The least soluble of the blood proteins is fibrinogen. The plasma is almost freed of it by half-saturation with sodium chloride, or with a small amount of acetic acid. It is precipitated as fibrin in the process of blood coagulation (see page 99), and is estimated by weighing the amount of fibrin which it produces.

Serum Globulin and Serum Albumin

Globulins are ordinarily defined as being insoluble in distilled water, and albumins as being soluble. It is, however, impossible to separate serum globulin and albumin satisfactorily in this manner. The globulin obtained by dialysis can be returned to solution by the addition of

a suitable amount of water, which makes the salt adherent to the precipitate a weak saline solution. In neutral or acid solutions it is coagulated by heat at about 75° C. But it does not act as an individual protein, since a portion of it is precipitated by dialysis or by carbon dioxide. Probably serum globulin really consists of two or more proteins.

The serum albumin remaining in solution after saturation with ammonium sulphate likewise does not represent a chemical entity. It is possible by carefully heating the solution of serum albumin to distinguish three separate coagulation temperatures. This fact has been interpreted as meaning that the serum albumin consists of at least three closely related proteins.

Since *the refractive index of the blood* depends primarily upon the amount of protein present, it has been taken as a means of determining variations in the concentration of the proteins. It has been found that the concentration of the blood proteins varies somewhat; during exercise it is increased probably because of the taking up of water by the tissues, and during profuse bleeding it is diminished because large amounts of fluid are being added to the blood from the lymph, which is relatively poor in proteins. The ingestion of considerable amounts of salts has been found to reduce the concentration of the blood proteins for a short time. In pathologic conditions, as in diabetes, when rapid changes in the body weight due to alterations in the diet are occurring, changes in the fluid content of the blood are often observed. Likewise in edema caused by faulty renal function, there may be a retention of fluid in the blood before there is any indication of edema. The hydremic condition of the blood can therefore be considered as a useful diagnostic aid in determining the water metabolism.

The relative concentration of the proteins of the blood is also of some interest, especially since in some diseases a considerable amount of blood protein is lost. By refractrometric methods it is possible to separate the globulin and albumin fractions. Normally the total proteins range between 6.7 and 8.7 per cent, of which the albumins lie between 4.95 and 7.7 per cent, and the globulins between 1 and 2.54 per cent. In some diseases, as in chronic nephritis, pneumonia, and syphilis, the total proteins of the blood are decreased and the relative amount of serum globulin is increased On the other hand, in many mild infections and chronic septic conditions the globulin fraction may be increased with no change occurring in the total protein content.[3]

Our knowledge of the *origin and the function of the blood proteins* is quite unsatisfactory. Previous to the discovery of amino acids, the building stones of the proteins, in the blood it was thought that the nitrogenous nutrients were converted somehow into blood proteins dur-

ing or immediately following their absorption from the alimentary canal, and that the tissue cells were nourished from this common protein. It is now known that the amino acids are not immediately synthetized into blood proteins after their absorption from the digestive system. The blood proteins are radically different from the tissue proteins. Substances which retard or accelerate nitrogen metabolism do not alter the relationship existing between the protein bodies of the blood. This fact indicates that the serum proteins have a function quite independent of the nitrogenous metabolism of the body. They undoubtedly maintain the viscosity of the blood and assist in preserving its neutrality. Attempts to localize the site of formation of the blood proteins have not been successful. There is some evidence that fibrinogen is formed for the most part in the tissues of the splanchnic area (liver). It is quite possible that the blood forms its own proteins, just as do other tissues, from the amino acids it contains.

THE FERMENTS AND ANTIFERMENTS OF THE BLOOD

The blood plasma contains many of the ferments present in the tissues. The nature of these ferments has been the subject of many investigations in recent years, primarily because it has been found that they are intimately connected with the problems of immunity.

Among the ferments the following have been demonstrated in the blood:

Proteases are probably present normally in the human blood serum in small amounts, but they are found in large amounts in the white blood corpuscles. A protein foreign to the body if injected into the blood ordinarily produces no untoward symptoms, but a second injection following the first by some days will produce symptoms of poisoning known as *anaphylaxis*. This fact has led to the assumption that the injection of any foreign protein into the blood promptly leads to the appearance therein of specific proteolytic enzymes which will digest the strange protein into its derivatives, which are poisonous. This power of the body to produce specific proteases has been the subject of much research and debate, and Aberhalden proposed a test for pregnancy, for cancer, and for other conditions in which he made use of this phenomenon. He believes the presence of placenta or tumor tissue to cause the presence of proteins that bring about the production of specific ferments whose duty it is to rid the system of these substances. Other investigators fail to find the specificity in proteolytic action claimed by Abderhalden, and believe that proteolytic ferments which are capable of digesting foreign proteins are absorbed from the alimentary canal

from the digestive juices (Boldyreff). Some investigators fail to confirm the claim that the proteolytic activity of the blood serum is increased under the above conditions.

Blood contains an antiferment known as *antitrypsin.* This can be removed from the blood serum by several substances, among which are kaolin, colloidal iron and starch. Serum thus treated shows strong proteolytic activity and autodigestion will occur. In this case there can be no question of the specific origin of proteases. Abderhalden believes that the ferments of the blood of the pregnant woman are able to digest the placental tissue. Human placental tissue has the ability of absorbing antitrypsin and it is very questionable as to whether the test proposed by Abderhalden is due to the new formation of ferments or to the removal of the antitrypsin and the action of the protease normally present in the blood.

Nuclein ferments are capable of decomposing nucleic acid and purins into the simpler bodies.

Lipases have been demonstrated in the blood.

Amylase.—The presence of starch-splitting ferments in the blood was first shown by Magendie in 1841, and later Bernard showed that glycogen or starch injected into a vein produced glycosuria. Since then it has been proved conclusively that diastatic enzymes are normally present in the blood and lymph. The source of these enzymes has given rise to much speculation. Some observers believe that they are derived from the amylopsin of the pancreatic secretion, while others believe that they are manufactured by the liver. Ligature of the pancreatic ducts is said to increase the amount of amylase, while removal of the pancreas may (Carlson and Luckhart) or may not (Schlesinger) increase the amylase of the blood. In some forms of experimental diabetes the amylase of the blood has been found increased, and this is the case in human diabetes (Myers and Killian). If this is true, a cause for the inability of the diabetic to store up glycogen is found. In impairment of renal function, there is usually an increase in the blood amylase and a decrease in the urine amylase. This has been suggested as being of diagnostic value.

The blood contains a feeble glycolytic enzyme capable of destroying glucose. It is claimed that this power is reduced in diabetics (Lepine).

Catalase is found in the blood and tissues generally. It has the power of liberating oxygen from hydrogen peroxide without any accompanying oxidation process. Its physiologic significance is not known. It is said that the amount of catalase is increased during excitement and exercise, and is decreased in conditions where the body's activity is lowered. Its determination is clinically unimportant at present.

CHAPTER XI

BLOOD: THE BLOOD CELL

BY R. G. PEARCE, B.A., M.D.

THE RED BLOOD CORPUSCLES, OR ERYTHROCYTES

The most prominent function of the blood is to carry oxygen to the tissues. It owes this property chiefly to the red blood cells which are present in large numbers (5,000,000 per c.mm. of blood). These cells are biconcave discs, having a diameter of about 7.7 μ. They are constructed out of a framework composed largely of lipoidal material, in the meshes of which is deposited a substance called hemoglobin, to which the remarkable oxygen-carrying power of the blood is due. Neither the manner by which the red cell carries its hemoglobin nor the intimate structure of the cell itself is accurately known. It is commonly believed that the hemoglobin is held enmeshed in a framework or stroma, or encased in the cell membrane. One thing is certain, however, that the union of hemoglobin with the stroma of the red cell is a fairly strong one, since mere fragmentation of the corpuscle fails to liberate the hemoglobin. The fact that the framework contains a large amount of lipoidal substances enables the corpuscles to maintain their shape and is responsible for their characteristic permeability.

Hemoglobin is a very complex substance belonging to the group of conjugated proteins. By chemical means it can be broken up into a simple globulin and a pigment hematin, containing iron. When completely saturated, oxygen is present in hemoglobin in the proportion of two atoms of oxygen to one atom of iron (Peters); or 401 c.c. of oxygen can be carried by hemoglobin containing one gram of iron, the molecular weight of the molecule being about 16.669, or some multiple thereof (Barcroft and Peters) (see also p. 397). At this figure the iron in the molecule would represent 0.34 per cent of the total weight of the molecule. The corpuscular surface area has been estimated to be 3200 square meters. There is therefore a very large surface available for the absorption of oxygen from the alveolar air, as the blood corpuscles pass in single file through the capillaries of the lungs.

Since the amount of oxygen which the blood can carry depends upon its hemoglobin content, it is of some importance clinically to have

methods of determining the approximate amount present. The amount of hemoglobin present in a quantity of blood is usually determined colorimetrically by comparing the color of the blood with standard colors which correspond to known strengths of hemoglobin. In normal persons the amount of hemoglobin varies greatly at different ages, and in order to determine whether or not a given blood contains more or less hemoglobin than normal, it is imperative to consider the age. The greatest variations occur between birth and the sixteenth year. After the sixteenth year the blood in males usually contains a larger amount than that in females (Williamson[4]). Instruments used in determining the amount of hemoglobin should be standardized to give the value in grams hemoglobin per 100 c.c. of fluid.

The amount of hemoglobin which is present in each corpuscle in terms of normal is therefore of some clinical interest. This relation of the number of red cells to the amount of hemoglobin is known as the color index and is computed as follows: The average red count in man is 5,000,000 to the c.mm., and the average minimal amount of hemoglobin is taken as 13.88 grams in 100 c.c. of blood (= 80, Sahli; = 90, Miescher; = 86, Plesch; and 110, Tallquist methods). These relative values give a color index of one. The percentage of normal red cells divided by the percentage of normal hemoglobin present gives the color index.

The Origin of the Red Blood Cells

In fetal life the spleen and the liver are generally believed to be responsible for the formation of the red blood cells. In extrauterine life this function is taken over by the red bone marrow. In the primitive condition all red blood cells are supposed to be nucleated. In extrauterine life the nuclei of the red cells are lost, and nonnucleated forms are alone present in the blood stream. In fetal life and in certain pathologic conditions, the rate of blood formation is so rapid that some nucleated cells appear in the blood. The normal response of the body to a loss of red blood corpuscles consists in an increased activity of the blood-forming cells of the red bone marrow. It is not easy to follow the course of the regeneration of the red corpuscles or to discover the mechanism of their formation in the bone marrow, since this tissue presents a mixture of cells which are precursors of the varied corpuscles found in the blood and the identity of which can not be determined.

Recently new methods of staining blood for microscopic examination have allowed more detailed study to be made on the site and method of blood cell formation. When fresh unfixed blood is treated with solutions of various dyes, such as brilliant cresyl blue, polychrome

methylene blue or neutral red, an otherwise invisible structure appears in some cells in the form of coarse granular particles or threads, which give a reticulated appearance to the corpuscles. These reticulated cells are more abundant in infants' blood and in patients suffering with severe anemia or hemolytic jaundice than in normal blood, and may be taken as evidence of the youth of the red cell and not as a degenerative process. Since the number of the reticulated cells that are present in the blood is more or less directly proportional to the hemopoietic activities of the bone marrow, enumeration of the reticulated cells is of clinical importance in anemias. In conditions in which animals have been made plethoric by the transfusion of blood, it has been found that the number of reticulated cells is decreased; the bone marrow of these animals also shows a marked reduction in reticulated erythroblasts. The diminished rate of blood cell formation sometimes noted after blood transfusions may be explained by assuming that the stimulus which awakens the formation of red cells in the bone marrow is absent or made subnormal on the injection of red cells into the blood, and thus the formation of red cells is depressed. Small transfusions are therefore preferable to large ones in cases in which the rate of blood formation is greatly impaired. By means of living cultures of red bone marrow the different stages of the development of the normoblasts into true red corpuscles may be studied (Tower and Herm[5]). Some evidence has been gathered from such studies which points to the conclusion that in place of the red cells being cells which have lost their nucleus, as is the current teaching, they are rather cells which develop as a nuclear bud and escape into the circulation as true red cells. The nucleated red cell and the red nucleated corpuscle of the bird are the product of intranuclear activity and are morphologically identical.

Rates of Regeneration of Erythrocytes

Microscopic examination of the blood during rapid regeneration of red cells shows the presence of nucleated forms. Nucleated red cells in the blood have therefore been taken as an inevitable feature of rapid blood regeneration. The evidence upon which this belief depends, however, is hardly complete, since changes in the manner of red blood cell formation may be responsible for the nucleated forms. The red bone marrow is considered the seat of red cell formation, and it is true that an abnormal increase in the red bone marrow usually accompanies increased red cell formation. The nature of the stimulus which brings about the new formation of red cells is not understood. Oxygen want may be an important factor, since we find the presence of an abnormally large number of red cells in conditions where there is a scarcity of

oxygen in the inspired air, as in life at high altitudes, or a difficulty in its absorption through the lungs, as in congenital heart disease.

The red cells produced following hemorrhage and in simple anemia contain less than the normal amount of hemoglobin, but their shape and size are approximately normal, and few nucleated cells are present. In the regeneration of red cells which is found in pernicious anemia, we find the cells containing an unusually large amount of hemoglobin. The red cells in this disease have abnormal forms, many being large, with or without a nucleus, and containing basic staining granules. This type of blood cell formation is due to degenerative changes.

The Fate of the Erythrocytes

The length of life of the red blood cell is unknown. Estimates based upon the daily excretion of bile pigments are not reliable, since Hooper and Whipple have shown that the pigments, in part at least, arise from pigments which the liver has made in excess of its needs for the manufacture of hemoglobin, and which, not being needed, are excreted.[5] There is no question however that every erythrocyte sooner or later undergoes disintegration, a process formerly thought to be ushered in by the ingestion of the red blood cell by a phagocyte in the spleen or in a hemolymph gland, the hemoglobin of the disintegrated cell being set free and carried to the liver, where it is broken up into hematin, which the body stores for future use, and into bile pigments, which are excreted. Rous and Robertson[6] fail to find evidence that this process occurs in man to an extent sufficient to account for the normal destruction of the blood cells. However they have recently found another and unsuspected method for blood destruction in all animals thus far studied—namely, the disintegration of the blood cells by fragmentation while they are circulating, without loss of their hemoglobin. These fragmented cells are found most frequently in the spleen. They believe that the small ill-formed cells, known as microcytes and poikilocytes, observed in severe experimental anemias, are due not to the fact that they are produced by the bone marrow, but rather to the fact that the marrow in its anemic condition is not able to produce a resistant erythrocyte, and fragmentation therefore takes place too readily. A similar condition may exist in the severe anemias of man and account for the general high resistance of the red cells found in the blood of these patients, inasmuch as the weak cells are generally fragmented very soon after they are formed. Long ago Ehrlich stated that the microcytes and poikilocytes of anemia are the result of fragmentation of the cells in the circulating blood, but he believed that this fragmentation was a

purposeful division in order to increase the total surface of the red cells. The ultimate fate of the red cell fragments is not known. It is reasonable to suppose that the fragmented bits containing hemoglobin are carried to the liver, where the hemoglobin is transformed into hematin and bile pigments.

Hemolysis

Another method of red blood cell destruction, which, however, does not take place normally, is by hemolysis. The nature of the combination of the hemoglobin with the stroma of the red cell, as already remarked, is not definitely known. That it is not merely contained in a sac is shown by the fact that the cell may be cut into bits without the hemoglobin being set free. In some manner the hemoglobin is chemically bound with the stroma of the red cell, from which it can be freed by a number of physicochemical and chemical agents. This process is known as *hemolysis,* and the substances which bring it about are known as hemolytic agents. The manner in which these agents effect the release of hemoglobin from the blood is quite varied.

If the osmotic pressure of the plasma is lowered by dilution, the pressure within the corpuscle remains high, and water is absorbed by the cell. If this absorption is sufficient, the cell ruptures and the hemoglobin is discharged. For this reason it is necessary in diluting the blood to use solutions of salt having an osmotic pressure equal to that of the blood to protect the red cell from hemolysis. This is obtained by using a 0.9 per cent solution of sodium chloride. Better results are had, however, by using either Ringer's solution (0.9 per cent NaCl, 0.026 per cent $CaCl_2$, and 0.03 per cent KCl) or Locke's solution (0.9 per cent NaCl, 0.024 per cent $CaCl_2$, 0.042 per cent KCl, 0.01-0.03 per cent $NaHCO_2$ and 0.1 per cent glucose).

In normal corpuscles hemolysis occurs to a small extent in solutions containing about 0.42 per cent of sodium chloride. In certain diseases the fragility of the corpuscles may be increased (Butler[7]).

The membrane and stroma of the erythrocyte contain lipoidal material which is soluble in alcohol, ether, fatty acids, and bile salts. Addition of these agents to the blood brings about hemolysis, presumably by dissolving the lipoidal material present. The hemolysis which occurs with saponin is similar in type, since saponins combine with lipoids, the compound being soluble in water.

The hemolytic properties of serum, whether they are found to be normally present when the bloods of certain animals are mixed or to be produced artificially by the injection of foreign red cells, furnish a subject of great interest from the standpoint both of immunology and

of clinical medicine. The hemolytic serum produced by the injection of foreign corpuscles owes its activity to two substances. The one called the *amboceptor*, or immune body, is specific against the type of cell injected and is increased during immunization. The second body is the complement; it is nonspecific, and is not increased during immunization. Complement is destroyed by heating the serum for one hour at 55° C., leaving the amboceptor alone present. Corpuscles placed in such serum are not hemolyzed until complement either from fresh immune or from nonimmune serum is added.

The serum of animals possessing natural hemolytic properties towards the corpuscles of other animals likewise owes its effect to the joint action of amboceptors and complement.

Ordinarily the serum from animals of one species does not exhibit hemolytic properties to blood from another animal of the same species. In unusual cases, however, the serum of an animal will produce hemolysis of the corpuscles of an animal of the same species. Such sera are said to possess isohemolysins. The fact is of great importance in the transfusion of blood from one individual to another.

The cause of the acute hemolysis which occurs in the disease paroxysmal hemoglobinuria is not known. It is probably due to the presence of a hemolytic substance which unites with the blood corpuscles at temperatures below the normal body temperature, since the attack follows exposure to cold, and blood from patients subject to the condition may be hemolyzed *in vitro* by cooling and subsequently heating it.

LEUCOCYTES

There are a number of varieties of white cells in the blood. These are differentiated from one another by their shape, staining properties, and the granules in their protoplasm. We may divide them into two main groups—nongranular mononuclear cells and granular polynuclear cells.

The nongranular mononuclear cells are termed *lymphocytes*. Two varieties are differentiated, the small and the large.

The small mononuclear leucocyte makes up from 23 to 28 per cent of the total leucocytes and the large mononuclear, from 2 to 4 per cent.

The polynuclear leucocytes are divided into three groups according to whether their granules stain with basic, neutral or acid stains. The leucocytes that stain with basic dyes, or the basophile cells, are very few, making up less than one per cent of the total count. Likewise the acid-staining granular cells, acidophile, are few, comprising from 2 to 4 per cent of the total count. The most numerous are the neutrophiles,

or the polynuclear leucocytes, with neutral-staining granules. These comprise from 65 to 75 per cent of the total count.

Another type of white cell is known as the transitional cell, because it was supposed to represent an intermediate form between the mono- and polynuclear cells. Probably such transitions do not occur, and the transitional leucocyte is related to the mononuclear cells.

The polynuclear cells originate in the bone marrow, and for this reason have been termed myeloid cells. They develop from cells in the bone marrow termed myeloblasts, which are nongranular and contain a large nucleus. In the course of development the characteristic granules appear, and the nucleus remains round and later becomes lobulated. These intermediate forms are called myelocytes. The mononuclear cells originate in the lymphatic tissues of the body.

The leucocytes possess the ability to make ameboid movement and to ingest foreign particles which may be presented to them. On account of this latter ability they are commonly called phagocytes. In the process of inflammation the leucocytes assemble at the spot which is the seat of the injury or infection, and remove the foreign organism or necrotic tissue by ingesting and digesting it.

It is not definitely known whether or not the lymphocytes function as phagocytes. Other functions besides those as phagocytes have been ascribed to the white cells, but they are not universally accepted. The number of leucocytes in the blood is subject to considerable variation. They normally number between 6,000 and 8,000 per c.mm. At the height of digestion and after strenuous exercise there is usually a small increase, and under pathologic conditions, especially in infectious diseases, this becomes quite marked. Some infections increase the polymorphonuclear cells, while others add to the lymphocytes. The factors governing the type of increase are not fully known, nor are the functions of the various forms differentiated.

The Blood Platelets

These are small oval particles about 3 μ in diameter, which are found in large numbers (250,000 to the c.mm.) in the blood. They are supposed to be formed from particles of protoplasm which are pinched off from the large blood cells in the bone marrow. Their biological and chemical properties are not understood. They probably play a very important role in the coagulation of the blood (see page 103).

CHAPTER XII

BLOOD: BLOOD CLOTTING

On leaving the blood vessels, the blood clots so as to form a plug, which assists in preventing further hemorrhage. The clotting must therefore be considered as a protective mechanism against excessive draining of blood out of the organism. When the wounded vessels are small, the clotting, along with constriction of the damaged vessels and the formation in them of thrombi containing large numbers of platelets, serves to effect complete stoppage of the hemorrhage even though the blood pressure may not have become materially reduced. The greater loss of blood from larger vessels causes the arterial pressure to fall, and this enables the clot to stiffen and seal the wound before the pressure again rises. When the clotting power of the blood is subnormal, life is endangered by even trivial wounds; under these conditions the smallest surface scratch may continue to bleed excessively in spite of whatever local treatment is applied. The most extreme degree of this condition occurs in hemophilia, a disease which is characterized by a most interesting family history—namely, that although it affects only certain of the male members of a family, yet it is transmitted from generation to generation by the female side alone. The disease has existed in certain of the royal families of Europe for many generations, which has made it possible by consulting the genealogic trees to demonstrate the infallibility of this law of inheritance.

The clotting of the blood is also either depressed or increased in a variety of physiologic and pathologic conditions. We shall, however, defer further consideration of these until we have learned something of the nature of the factors which are responsible for the process itself.

The Visible Changes in the Blood During Clotting

In a few minutes after it leaves the blood vessels, the blood forms a jelly-like clot, which adheres to the walls of the container in which the blood is collected and soon becomes so solid that the vessel may be inverted without spilling any of the blood. Clotting is now said to be complete. The clot soon begins to contract, and as it does so, drops of clear fluid or serum become expressed and float on the surface of the

clot or collect between it and the walls of the container, so that after some time the clot breaks away from the container and comes to float in the serum. The latter may be perfectly clear, but usually is more or less opalescent, partly because of the presence of fat, and partly because of leucocytes which have migrated out of the clot on account of their power of diapedesis.

If a drop of freshly shed blood is examined under the microscope, it will be observed that the first step in clotting consists in the formation of fine threads radiating from foci, which are undoubtedly the blood platelets. The fine threads are called *fibrin*. They multiply rapidly, so as to form an interlacing meshwork which entangles the red blood corpuscles and leucocytes. By the use of the ultramicroscope (page 52), Howell[1] and others have observed that the fibrin (produced by adding thrombin to oxalated plasma) is really deposited in the form of fine crystalline needles—"fibrin needles"—which become packed together as they increase rapidly in numbers. Although the process of clotting consists therefore in the conversion of a hydrosol into a hydrogel (see page 60), it is a unique process; a solution of the blood protein which is responsible for the formation of the fibrin (fibrinogen) may, like other colloidal solutions, be precipitated in a variety of ways, but it is only when the conditions are favorable for blood clotting that fibrin needles, and therefore fibrin threads, are formed. The blood of invertebrates forms a structureless gel when it clots (Howell).

Methods of Retarding Clotting of Drawn Blood

To understand the nature of the clotting process and the factors that are responsible for its occurrence, it is advantageous to simplify the conditions somewhat by getting rid of the red corpuscles and most of the other formed elements of the blood and then using the fluid in which these are suspended in living blood—namely, *the plasma*. This separation of blood into corpuscles and plasma is readily effected either by sedimentation or by centrifuging after measures have been taken to inhibit or greatly delay the clotting process. The methods used for this purpose are numerous. A few of the most important are as follows: (1) Keeping the blood at a temperature very slightly above freezing point. This method is, however, not very effective unless the blood is immediately received into narrow vessels placed in ice and the temperature kept most strictly at the low level. In the case of horses' blood and other slowly clotting bloods, the method succeeds without these precautions. (2) Receiving the blood through a strictly clean and smooth cannula, coated with a layer of paraffin or vaseline, into a vessel similarly coated. This method is of practical importance when it is necessary to

transfuse blood without making a vessel-to-vessel anastomosis. (3) Mixing the blood with chemicals that are capable of removing the calcium from solution. Such reagents are potassium or sodium oxalate (in a concentration of 0.1 per cent after mixing), and sodium fluoride and sodium citrate (2 per cent solution, with one part of the solution to four parts of blood). (4) Mixing the blood with certain neutral salts, particularly the sulphates of sodium and magnesium (one part of 27 per cent solution of magnesium sulphate mixed with four parts of blood). Blood thus treated is known as "salted blood," and the plasma separated by centrifuging, as "salted plasma." Clotting is readily induced by adding water to the salted blood or plasma, and in this way diminishing the concentration of the salts. (5) The addition to blood of one of a class of substances known as antithrombins. Leech extract or the purified substance separated from it, known under the trade name of "hirudin," and substances present in blood removed from animals after they have been injected with peptone solutions, are examples.

The methods which have just been described are those applied to blood after it has left the blood vessels. Another interesting group of anticoagulants prevent clotting *only when injected into the blood vessels of the living animal.* The most powerful example of this group is snake venom, certain varieties of which can prevent clotting in the dosage of $\frac{1}{100}$ of a milligram for each kilogram of body weight. Similar but much less potent effects are produced by the injection of several proteolytic enzymes, but most attention has been paid to the effect of commercial peptone injected in solution intravenously in the proportion of 0.3 gram to each kilogram of body weight. Blood subsequently removed up to about half an hour or more does not clot, and as we have already seen, if added to blood from another animal, materially retards clotting. This group of *intra vitam* anticoagulants is particularly interesting, since none of the substances belonging to it is capable of preventing clotting of blood when mixed with this after it has been shed. Their action therefore obviously depends on the production of some substance in the body, probably, as we shall see later, in the liver, since they fail to act after the removal of this organ from the circulation (see page 111).

The time of clotting varies greatly according to the conditions under which the blood is collected and the animal from which it is derived. Human blood, for example, received into a test tube from a puncture through the skin may clot at any time within three or ten minutes, five minutes being taken as an average time for blood kept at a temperature of about 20° C. This time may be considerably shortened by increasing the extent of foreign material with which the blood comes into contact, and more particularly by whipping the blood with a bunch of twigs or

wires. In this latter case, however, the clot does not form in the usual manner, but the fine threads of fibrin collect on the twigs or wires, leaving behind the blood serum with the corpuscles still suspended in it. The fibrin removed in this way may then be washed free of adherent serum. The serum and corpuscles now form *defibrinated blood,* which is used for many physiologic purposes. Clotting is also greatly accelerated by allowing the blood to flow over exposed tissues. Something is evidently added to it from the tissues which accelerates the clotting process, this influence being particularly marked in the case of blood of the lower vertebrates. When the blood of the bird, for example, is received through a cannula inserted directly into a vessel with as little injury to the walls as possible, it very slowly clots if at all, but soon does so if the blood is allowed to come into contact with excoriated tissues, or if it is mixed with tissue extract, such as that of muscle. Clotting is considerably accelerated by warming the blood. The application of a cloth or tampon well wrung out with hot physiological saline to a wounded surface is a most efficient means of allaying hemorrhage from vessels too small to ligate.

The Nature of the Clotting Process

Plasma obtained by centrifuging blood that has been prevented from clotting by one of the foregoing methods can be made to clot by removing the inhibiting influence; for example, in cooled plasma by warming the blood to room temperature, in salted plasma by diluting it with at least an equal volume of water, and in decalcified plasma by adding a sufficient amount of soluble calcium salts to combine with all the added oxalate and leave a small trace of calcium salts in excess.

The first question concerns *the source of the fibrin,* and the answer to it is furnished by comparing the composition of bood plasma with that of serum. Though both of these fluids contain the proteins, albumin and globulin, in approximately the same concentrations, the plasma also contains another protein not unlike globulin in most of its reactions, but distinguished from typical globulin in that it is precipitated by half-saturation with sodium chloride, in which typical globulin is soluble, and is more readily coagulated by heat. To produce half-saturation of the plasma with sodium chloride, equal volumes of plasma and saturated sodium-chloride solution are mixed together. The precipitate of ***fibrinogen,*** as the substance is called, is then collected at the bottom of the tube by centrifuging and is washed several times by decantation with half-saturated sodium-chloride solution. The washed precipitate, dissolved in weak saline solution (preferably containing a trace of bicarbonate), will then be found to clot under certain conditions.

The next question concerns *the nature of the conditions that cause the fibrinogen to clot.* When a fibrinogen solution is mixed with a few drops of blood serum, a clot usually forms, which however is not the case when plasma is added or when the serum is heated before adding it. Because a small quantity of serum is capable of causing the clotting of a large quantity of fibfinogen solution or plasma, it is supposed that the active substance present in it is of the nature of a ferment—*fibrin ferment* or *thrombin.* It must be pointed out, however, that there is considerable doubt whether this active body is really of the nature of a ferment or enzyme. For example, although heated serum does not cause clotting, thrombin, prepared from serum by the method about to be described, in the absence of inorganic salts can withstand even a boiling temperature. Moreover, true enzymes are characterized by the fact that, like other catalytic agents, a very minute quantity can effect a change in an indefinite amount of substance without the enzyme becoming used up in the process (page 72). When thrombin is allowed to act upon a fibrinogen solution, on the other hand, it is said that only a fixed amount of fibrin can be formed when a small amount of thrombin is added. Neither does this amount increase when the time of reaction is prolonged.

Whatever may be the significance of the foregoing facts, it is important to know that the clotting substance, thrombin, can be isolated from blood serum in a tolerably pure condition. For this purpose blood serum is allowed to stand under a large volume of alcohol for a week or two; the precipitate is then collected and rubbed up with water, which extracts the thrombin from it, leaving the serum protein in a coagulated state. The resulting watery solution of thrombin may be further precipitated by alcohol, the precipitate washed in alcohol and redissolved in water, yielding ultimately a solution which exhibits very marked coagulating powers when added to plasma or fibrinogen solution. Thrombin shows most of the protein reactions but it is not coagulated by heat. As would be expected, a considerable quantity of thrombin remains adherent to the fibrin formed in the process of clotting, and Howell[8] describes a very useful method by which it can be separated from fibrin and preserved in a dry condition. Briefly stated, this method consists in allowing washed fibrin to stand overnight under eight per cent sodium-chloride solution, which dissolves the thrombin. The resulting extract is then mixed with an equal volume of acetone, which throws down a precipitate containing the thrombin. To preserve it, the precipitate is collected on a number of small filter papers, which are subsequently opened out and dried by exposure to a current of cold air before an electric fan. When the thrombin solutions are desired, the dried precipitates are extracted with a little water.

Thrombin does not exist in blood plasma, for if a clean and paraffined glass tube is inserted into an artery and the blood collected under alcohol, the precipitate after standing a few weeks will yield no thrombin when triturated with water. Quite clearly, therefore, the thrombin is produced at the time the blood clots, and the question arises, What is it produced from? It will be remembered that, when the blood is examined under the microscope during the clotting process, the fibrin threads are seen to start from foci which correspond to the blood platelets. It would appear therefore that the thrombin must be derived from some substance that is shed forth from the platelets during the disintegration which they undergo shortly after the blood is shed. The substance is called *prothrombin*. The platelets or their precursors, the megacaryocytes of red bone marrow, are probably not its only source, for clotting may occur in the complete absence of platelets, when it appears to come from the leucocytes. Prothrombin appears plentifully in the fluid used to perfuse red bone marrow outside the body (Drinker and Drinker[9]).

To sum up what we have so far learned, it may be stated that the process of clotting starts with the disintegration of blood platelets and probably of leucocytes, as a result of which there is shed forth into the plasma a substance called prothrombin, which immediately afterward becomes activated or converted into thrombin. The thrombin then attacks a protein present in plasma called *fibrinogen*, producing from it in thread-like form the insoluble protein, fibrin. But this does not complete the history, for at least two other important factors come into play; the one is the presence of soluble calcium salts, and the other that of peculiar substances derived from the tissues outside the blood vessels and called thromboplastic substances or thromboplastin (Howell). We must now consider the action of these two factors.

The Influence of Calcium Salts.—As already explained, the proof that soluble calcium salts are necessary for clotting is furnished by the observation that the process is entirely prevented when the freshly drawn blood is mixed with soluble oxalate. To this proof, however, objection might be made on the score that the oxalate *per se* inhibited the clotting. That such is not the case is indicated by the fact that, if the oxalated blood or plasma is dialyzed against physiologic saline solution till all the soluble oxalate has been removed from it, clotting is still absent but immediately supervenes if some soluble calcium salts are added. The question arises as to how the calcium ion acts. Two possibilities exist: (1) that it is concerned in the conversion of fibrinogen to fibrin, and (2) that it is necessary for converting prothrombin into thrombin. It can quite readily be shown that it is by the second of these processes

that the calcium acts; for example, clotting occurs when purified thrombin is added to dialyzed oxalate blood or plasma or to a pure solution of fibrinogen. Citrates prevent clotting by forming calcium citrate, which although soluble does not ionize in solution. It is the free calcium ions that are important. The action of the fluoride is somewhat mysterious, for it has been found that to produce clotting in fluoride plasma the simple addition of calcium chloride will not suffice; thrombin itself must be added as well. Some authors assert, however, that if the calcium chloride is added cautiously to "fluoride" blood, it will induce clotting (Rettger). In any case it appears that the fluoride does something more than precipitate the calcium; possibly it prevents the breaking up of platelets and leucocytes.

The Influence of the Tissues.—As already stated, when slowly clotting blood, like that of a bird, is collected through a sterile glass tube into a thoroughly clean vessel and immediately centrifuged, the plasma will often remain indefinitely unclotted. If an extract of some tissue, such as muscle, is added, however, the plasma immediately clots. To a much less degree, the same phenomenon is exhibited by mammalian plasma when it is collected in a similar manner. From these observations the conclusions may be drawn that the tissues furnish some substance assisting in the clotting process, and that this substance is also formed from certain elements present in mammalian but not present in avian blood. The absence of platelets from the latter blood suggests that these must be the source of the activating substance in mammalian blood. It is plain that this tissue factor in clotting is of importance in hastening the process when an animal is wounded.

Before attempting to formulate an hypothesis that will explain the process of clotting, it is necessary to call attention to one other important fact. This refers to the presence in blood of a substance that prevents clotting and is hence called *antithrombin*. Antithrombin is present in normal blood, for a given specimen of pure fibrinogen will clot less rapidly when mixed with serum to which some oxalated plasma has been added than with an equal amount of the same serum correspondingly diluted with a solution of soluble oxalate. A striking increase in the concentration of antithrombin in blood can be brought about by rapidly injecting a solution of commercial peptone into the blood vessels fifteen to thirty minutes before bleeding. The peptonized blood or plasma will remain fluid for many hours, if not indefinitely. That the failure of this blood to clot depends on the presence of some anticlotting substance, and not upon the absence of one of the necessary clotting substances (fibrinogen, thrombin, etc.), is evidenced by the fact that the addition of some of it to a mixture of thrombin and fibrinogen inhibits

the coagulation, which it does not do, however, if it is first of all heated to 80° C. and filtered free of the coagulated protein. Moreover, the antagonistic action is quantitative in the sense that a fixed amount of the peptone-plasma inhibits the action of a fixed amount of thrombin. The source of antithrombin in the body appears to be mainly at least the liver, for it has been found: (1) that peptone injection into an animal from which the liver has been removed does not cause antithrombin to be formed (Denney and Minot);[10] (2) that peptone injections into the portal vein cause antithrombin to appear in the blood much more rapidly than when the injection is made into a systemic vessel; and (3) that, when the liver is perfused outside the body with a perfusion fluid containing peptone, antithrombin accumulates in the perfusion fluid.

A fluid containing a high concentration of antithrombin is secreted by the so-called salivary gland at the head end of the leech. The function of the fluid is to prevent clotting of the blood, so that the animal may continue to suck it without interference by clotting. After applying leeches for medicinal purposes it is therefore necessary to wash the wound thoroughly with water so that all traces of the antithrombin may be removed; otherwise the bleeding may continue for a considerable time. Practical use is made of this effect of the leech to prevent clotting of blood outside the body, or it may be used to inhibit coagulation *intra vitam* in experiments where clotting would otherwise interfere with their progress; for example, in crossed circulation experiments (page 365) and in experiments in vividiffusion (page 607). For such purposes the leech head is cut off and extracted either with saline or by treatment with chloroform, which removes other proteins from the saline solution leaving a strong antithrombin, known under the trade name of "hirudin." At temperatures about that of the body the action of antithrombin is greatly augmented. In animals like the mammals in which the content of antithrombin is small, this may be important in maintaining the fluidity of the blood (Howell). Blood containing antithrombin can be made to clot by the addition of thrombin, and therefore of blood serum.

CHAPTER XIII

BLOOD: BLOOD CLOTTING (Cont'd)

THEORIES OF BLOOD CLOTTING

Attempts to link all the foregoing facts together in the form of a simple theory have not so far been entirely successful. All agree that the fibrin is derived from fibrinogen by the action of thrombin, the points in dispute being those which concern the origin of the thrombin and the mode of action of the calcium and thromboplastic substances. The theory most widely accepted in Europe is that of Morawitz, according to which the thrombin exists in living blood in an inactive state called thrombogen (prothrombin), which becomes converted into thrombin by the simultaneous action on it of soluble calcium salts and of thromboplastic substances furnished by the tissue cells in general and by the cellular elements of the blood platelets and leucocytes. According to this view the thromboplastic substance, aided by the presence of calcium ions, converts thrombogen (prothrombin) to thrombin. It acts therefore as a kinase and is called thrombokinase. The fundamental fact of this theory, then, is that kinase is necessary for the union of the calcium with prothrombin—a fact, however, which is challenged by Howell, who states that prothrombin may be converted to thrombin by the action of calcium ions alone. This investigator believes that the thromboplastic substance acts not as a kinase but because it neutralizes antithrombin, which is constantly present in the blood, and the function of which is to prevent the calcium from uniting with the prothrombin to form thrombin. Howell's theory in his own words is as follows: "In the circulating blood we find as constant constituents fibrinogen, prothrombin, calcium salts and antithrombin. The last named substance holds the prothrombin in combination and thus prevents its conversion or activation to thrombin. When the blood is shed, the disintegration of the corpuscles (platelets) furnishes material (thromboplastin) which combines with the antithrombin and" at the same time liberates more "prothrombin; the latter is then activated by the calcium and acts on the fibrinogen." Antithrombin can also prevent the action of thrombin on fibrinogen. As already pointed out, the thromboplastin can be derived from the blood itself in the mammals, but only from the tissues in the lower vertebrates. It is interesting to note that the thromboplastin

can be extracted from the tissues by fat-solvents, and that it appears to belong to the class of phosphatids, being indeed closely related to, if not identical with, kephalin (Howell).

Intravascular Clotting

The practical application of the theory of blood clotting concerns the manner in which the blood is maintained in a fluid condition in the blood vessels, and the disturbance of this function causing *intravascular clotting*. According to the one theory, the blood is maintained fluid by the absence from it of any considerable quantity of kinase, and according to the other, by the presence in it of an amount of antithrombin sufficient to prevent the union of calcium with prothrombin. The fluidity is maintained even when large amounts of thrombin or of blood serum, which contains this substance, are injected into the living animal. We can best explain the immunity of the blood to the action of thrombin under these circumstances as being due to the instantaneous appearance in it of antithrombin in amounts sufficient to prevent the action of thrombin on fibrinogen, for, as stated above, it is claimed by Howell that antithrombin has this influence as well as that of preventing the conversion of prothrombin into thrombin.

Intravascular clotting may be brought about by a variety of means: (1) Mechanical damage to the lining of the blood vessels; after the application of a ligature, for example, the damaged endothelium is soon covered by a clot, which gradually becomes firmer and firmer, and may spread up the vessel to the next branch. (2) The presence of foreign substances in the blood. Emboli, for example, are apt to cause clots to form at the places where they stick, namely, in the smaller vessels. Clotting is also a frequent occurrence when there are local dilatations of the cardiovascular tube, and it may occur under imperfectly understood conditions causing the condition known as *thrombosis*. (3) An interesting variety of intravascular clotting is that caused by the intravenous injection of saline extracts of cell-rich tissues, such as the thymus, lymph glands or testes (Wooldridge). By precipitation with acetic acid and digestion with peptone, a residue can be obtained from these extracts which, when dissolved in alkali, has a very pronounced intravascular clotting effect. Since these precipitates are very rich in phosphorus, it is probable that they are of the nature of phosphoprotein (nucleoalbumin). Their action must depend on neutralization of antithrombin, according to Howell's theory, or because they serve as thrombokinases (according to Morawitz' theory).

As a matter of fact, however, the foregoing observation is not completely explained by either theory. If in place of making one injection

frequent injections of small amounts of the above material are made, instead of intravascular clotting, a delay in the coagulation time is likely to occur. Indeed, repeated injections of small amounts may entirely remove the clotting power of the blood. The readiness with which this so-called "negative phase" appears, seems to depend on the nutritive condition of the animal at the time of injection. If a large dose is injected into a fasting dog, for example, thrombosis is confined to the portal area, whereas if it is injected into a recently fed animal, the thrombosis is universal throughout the vascular system. The development of the negative phase is undoubtedly dependent upon some reaction on the part of the living cells of the organism, since it does not occur on the addition of similar substances to blood outside the body. The reaction is, indeed, akin to that by which immune bodies in general are produced. For example, a toxin injected in large amount has a certain toxic effect, but in repeated small doses with intervening intervals it leads to the production of an antitoxin. So with the substance in question; a large dose injected at one time causes a positive effect—clotting—but smaller doses frequently injected, the opposite effect—want of clotting. It is probable, as suggested by Starling, that more intensive study of the conditions causing intravascular clotting will throw considerable light on the general question of the production of immunity.

Measurement of the Clotting Time

To measure the clotting time of drawn samples of blood, several conditions must be observed. These have been tabulated by Addis[11] as follows:

1. The specimens of blood must always be obtained by exactly the same technic. It would introduce serious errors to compare the clotting time of one specimen of blood received from an incision of the skin (ear lobe) with that of another collected in a syringe by venipuncture.

2. The temperature conditions must always be the same. Probably 25° C. is the best temperature to use. Higher temperatures are unsuitable for two reasons: first, because during its collection the blood will have become cooled to about or below this point, and time would be consumed in raising it higher; and second, because the time of coagulation is more and more shortened for each degree that the temperature is raised, this acceleration becoming especially pronounced for temperatures above 25° C. Quite apart from the liability to incur errors incident to measurement of shorter periods of time, observations at higher temperatures necessitate most rigorous adherence to a fixed temperature of the water-bath. Temperatures much below 25° C. are unsuitable, be-

cause the clotting sets in gradually and it is difficult to tell precisely when it occurs.

3. The blood must always be collected in the same sort of vessel and come in contact with the same kind and amount of foreign material. To this it may be added that the receiving vessel must be scrupulously clean; any trace of old blood clot or of serum is especially to be guarded against.

4. The end point must be sharp. It is here that the greatest technical difficulties are met with in making precise measurements, and it is greatly to be desired that different investigators should adopt some uniform method. For experimental purposes the method of Cannon and Mendenhall[12] is no doubt the best, and it has the added advantage of giving a graphic record of the observations. The accompanying figure (Fig. 19) shows the principle of the method. The blood is received through a standard cannula (*C*) into a tube (*T*) 5 cm. long and of 5 mm.

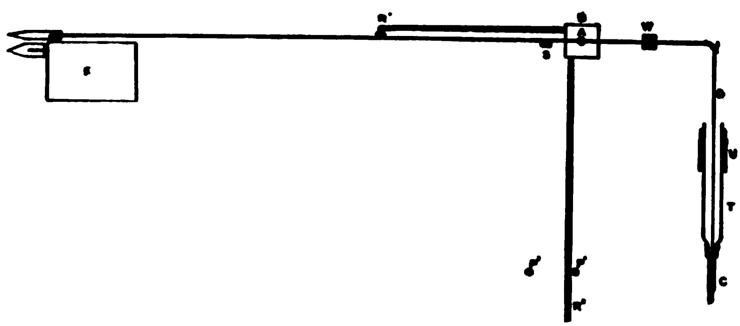


Fig. 19.—Diagram of the graphic coagulometer. The cannula at the right rested in a water bath not shown in this diagram. For further description see text. (From Cannon and Mendenhall.)*

internal diameter; and a loop (of 2 mm. diameter) at the end of a copper wire (*D*), which is 8 cm. long and 0.6 mm. in diameter, is allowed to fall gently into the blood at regular intervals. The upper end of the wire is articulated with the short arm of a light lever so counterpoised that when the stop (*R*), which ordinarily holds it in a horizontal position, is released, the wire, now having a net weight of 30 mg., falls on the blood in the tube. The long arm of the lever is provided with a writing point, which is made to inscribe its movements on a drum. So long as the blood is unclotted the loop sinks into it when the lever is released and a vertical line is traced, but whenever clotting occurs the loop sticks on the blood and the writing point does not rise.

For clinical purposes where blood collected in a syringe by venipuncture is used, the method of Howell[13] is most accurate. It consists in placing

*Am. Jour. Physiol., May 1, 1914, xxxiv, No. 2.

2 or 4 c.c. of the blood in a wide tube (of 21 mm. diameter) that has been cleaned by a bichromate-acid mixture. The period that elapses between the moment of the entry of fluid into the syringe and that at which the clot has become firm enough so that the tube can be inverted without spilling any blood, is taken as the clotting time. Since the blood does not come in contact with exposed tissues, it takes from 20 to 60 minutes to clot by this method.

For routine clinical examination of blood taken from a skin wound Brodie and Russel's method[14] is most satisfactory. This consists in principle in observing a drop of blood, under the low power of the microscope, while a fine current of air is gently blown against it at regular intervals in a tangential direction. Until clotting sets in, the individual corpuscles move freely in a circular direction, but as soon as clotting begins they move in masses which soon tend to become fixed so that, although they move somewhat when the air impinges on them, they immediately return to their original position when the current is discontinued. When clotting is complete, the air current merely

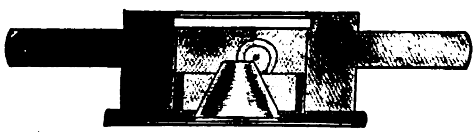


Fig. 20.—Coagulometer. The drop of blood is placed on the lower end of the glass cone and the air stream is directed against it from the side tube shown by the black dot. The apparatus is placed on the stage of the microscope and the drop observed by the low power.

presses on the corpuscles at one point. By this method the clotting time averages five minutes. A convenient apparatus for this method is that of Boggs, which is shown in Fig. 20. It consists of a truncated cone of glass, projecting into a moist chamber provided with a tube on the side so arranged that when air is blown into the chamber, it strikes the drop of blood placed on the end of the cone tangentially.

Blood Clotting in Certain Physiologic Conditions

Besides the experimental conditions already enumerated as changing the clotting time in the blood of laboratory animals, special mention must be made of the influence of epinephrine injections, of conditions supposed to cause a hypersecretion of this hormone, of the emotions, and of hemorrhage.

Epinephrine added to drawn blood does not affect the clotting time, but if small amounts are injected intravenously or even subcutaneously, a marked decrease occurs (Cannon and Gray; cf. Cannon, loc. cit.). Larger injections may have the opposite effect, and intermediate amounts

may cause at first a prolongation and later a shortening of the time. These results with larger doses are related to Howell's observation that repeated doses of relatively large amounts of epinephrine in dogs may so greatly retard coagulation as to make the animals practically hemophilic. It was further found by Cannon and his coworkers that epinephrine does not influence the clotting time when injected into animals from which the abdominal viscera have been removed from the circulation by ligation of the inferior vena cava and the abdominal aorta. In the light of the influence which destruction of liver cells (by phosphorus, chloroform, etc.) is known to have in lengthening clotting time, it would seem reasonable to conclude that it must be through this organ that epinephrine develops its clotting effects.

Stimulation of the splanchnic nerves also shortens the clotting time, and it would appear that this action depends on the resulting hypersecretion of epinephrine (page 746), for it is not observed following stimulation of the nerves in animals from which the adrenal glands have been excised (Cannon and Mendenhall). The interesting suggestion is made by Cannon that the shorter clotting time observed in animals showing strong emotions of fright or fear may also be due to the hypersecretion of epinephrine which this worker believes accompanies such states.

Blood Clotting in Disease

With the factors concerned in the process so wrapped in mystery, it is not surprising that the underlying causes responsible for delayed or deficient clotting of blood in diseased conditions or for the formation of intravascular clots (thrombi) are little understood. According to Howell's theory of the nature of the process, which is the most satisfactory at the present time, abnormal clotting might be due to the following causes: (1) *A diminished amount of fibrinogen.* This occurs when the hepatic cells are greatly damaged, as in poisoning by chloroform or phosphorus and in such diseases as acute yellow atrophy and yellow fever. In many cases of chronic cirrhosis of the liver, as Whipple, etc.,[15] have shown, the blood also clots feebly because of deficient fibrinogen. It should be pointed out that it is not so much the clotting time that is increased in such cases, as the firmness or consistency of the clot that is affected.

2. *A deficiency in prothrombin.* In the condition known as "melena neonatorum," undoubted benefit is derived from intravenous injections of blood serum or by direct blood transfusions, probably because thrombin or prothrombin is thus furnished.

3. *A deficiency of thromboplastin.* Since this substance is derived from both blood cells and tissue cells, it does not seem likely that a deficiency

could ever occur. Certain observers, however—Morawitz, for example—lay great stress on this as an important factor in hemorrhagic diseases.

4. *An excess of antithrombin.* The undoubted increase in this substance that can be brought about experimentally by injecting hirudin or peptone into animals, has stimulated careful search for a similar increase in the blood in clinical conditions in which abnormal blood clotting is one of the symptoms (Whipple[16]). Antithrombin is said to be increased in septicemia, pneumonia, miliary tuberculosis, etc.

5. *A deficiency of calcium ions.* Although at one time it was supposed that this might be responsible for the feeble clotting in hemophilia, it has not been found, after very extensive trials, that the exhibition of Ca salts in any way relieves the condition. It is said, however, that the slow coagulation seen in obstructive jaundice is decidedly shortened by treatment with calcium salts.

One thing stands out prominently in connection with the whole problem, and that is the close relationship of the blood platelets to the clotting process. From these cells are derived, according to Howell, not only the prothrombin but also, as from other cells, thromboplastin. It is not surprising therefore to find that decided alterations in the platelet count occur in cases of faulty blood clotting, and that local accumulations of these elements within the blood vessels, produced by their clumping together or agglutinating, is followed by a formation of local clots, as in thrombosis.

Hemorrhagic Diseases

In many of the so-called hemorrhagic diseases (acute leucemia and aspastic anemia) and in the hemorrhagic varieties of diphtheria and smallpox, the platelet count drops from its normal of between 200,000 and 800,000 per cubic millimeter to well below 100,000, and indeed in these conditions it is frequently difficult to find any platelets. Samples of blood clot outside the body within the normal time, but the clot is soft and usually fails to retract in the normal manner. It is on account of this, rather than slow clotting that the hemorrhage continues, so that in appraising the gravity of the symptom it is best to measure not the clotting time but the time that it takes for bleeding to cease from a small skin wound, as in the lobe of the ear. This can be very accurately done by applying blotting paper at regular intervals to the puncture (Duke[17]).

The most interesting and at the same time the most mysterious of all conditions in which blood clotting is interfered with is *hemophilia.* The clotting time is longer than normal, but even after the clot forms, bleeding is likely to continue because the clots are very readily displaced. Both clotting time and bleeding time are increased. So far no change in the

clotting factors of the blood has been demonstrated in this disease; the corpuscles and the platelets are normal in numbers, fibrinogen and calcium salts are normal, and, as Howell has shown, there is no excess of antithrombin. One significant fact, however, is that the addition of thromboplastin or of its active ingredient, kephalin, greatly shortens the clotting time of the blood when it is removed by venipuncture. In agreement with this observation it has been found that hemophilic blood clots much more rapidly, indeed sometimes in the usual time, if it is allowed to flow over cut or damaged tissue and so become mixed with thromboplastin. These facts taken together would seem to indicate that the fault must lie in a deficiency in prothrombin, and since this is derived mainly from the platelets, which however are not decreased in number, we must further assume that these elements have undergone some qualitative change preventing their disintegration. An accompanying defect in their agglutinating properties would at the same time explain their failure in hemophilia to clump together at the site of the hemorrhage so as to block the smaller vessels with thrombi; hence the prolonged bleeding time even after clotting has occurred.

Thrombus Formation

The first formed portion of a thrombus is paler than those formed later, because it contains excessive numbers of platelets; and it seems clear that it is by agglutination of these into masses, which then stick in the blood vessels and by disintegrating shed forth prothrombin and thromboplastin, that the clotting starts. This platelet agglutination may result from stagnation in the bloodflow, or from roughening and damage to the vessel walls. Stagnation may be due either to failure of the circulation as a whole as in heart disease, or to local physical alterations in the vascular tube, setting up conditions in which eddy currents with stagnant pools of blood are formed, such as will occur at places where the vessels suddenly become wider, as in varicose veins, in aneurisms and at the sudden bend of large veins. The first formed (platelet) thrombus is followed by one of a darker color, which fills the vessel up to the next anastomotic branch. Similar stagnation may also follow the obstruction caused by lodgment of emboli in the smaller vessels (air, foreign bodies in fine suspension, bacteria, etc.). The thrombi in such cases are very small and occur particularly in the capillaries of the liver, spleen, and lungs. The small thrombi often serve as foci from which clotting spreads into the larger vessels, this being often encouraged by an increase in the coagulability of the blood. When the intima is inflamed, it is possible that excessive amounts of thromboplastin are produced and that this neutralizes the antithrombin in blood moving so slowly that it is not

replaced by fresh blood before clotting ensues, or it may be that substances derived from the inflamed tissue cause the platelets to agglutinate. The increased clotting often observed after the injection of hemolytic agencies (foreign sera, snake venom, etc.) may also be due to platelet agglutination. Like the thrombosis following embolism, the clotting occurs at first in the capillaries, the initial thrombi containing masses of platelets along with skeletons of blood corpuscles and cells from the blood-forming organs.

CHAPTER XIV

LYMPH FORMATION AND CIRCULATION

GENERAL CONSIDERATIONS

Lymphatics are modified veins. They grow from the veins in embryonic life as buds of endothelium, which are first visible in the human embryo in the sixth week of development. The earliest outgrowth occurs from the internal jugular vein, and the endothelial buds soon become hollow and join together, forming first a plexus and subsequently a sac, from which again lymphatic vessels made of endothelium grow out to invade the skin of the head, neck, thorax and arm, and partly the deep structures of the head. The sac is ultimately transformed into groups of lymph glands. At a later stage similar nodes develop from certain of the abdominal veins, forming a retroperitoneal sac, from which grow out the lymphatics of the abdominal and, to a certain extent, of the thoracic viscera. A similar pair of sacs also develops from the iliac veins supplying the lymphatics for the skin of the legs and abdominal walls. The retroperitoneal and iliac sacs then become connected with the jugular sac by means of the thoracic duct. In the embryo there are no valves in the lymphatic vessels, so that the whole system can be injected either from the thoracic duct or from the skin, showing clearly that the superficial and deep lymphatics are parts of one closed system of vessels.

Anatomists have succeeded in tracing the course of the lymphatics in many parts of the body. This knowledge is of great importance in connection with the spread of infections, etc. Lymphatics are abundant in the skin, the intestine, and connective tissues, but are absent from the muscle bundles, from the hepatic lobules (though present in the connective tissue between them), from the substance of the spleen, and from the central nervous system.

The lymphatics have the same functions as blood capillaries, namely, to absorb substances from the tissue spaces. There is some evidence to show that this absorption may be selective. When injections are made into the peritoneal cavity, the pathway of absorption may be either the blood vessels or the lymphatics, according to the nature of the substance injected. True solutions are absorbed by the blood, but granules

are taken up by special large cells showing phagocytic powers, and transferred to the lymphatics—for example, those of the diaphragm. A similar selective absorption is well known in the case of the villi of the intestine, where fat passes into the lacteals and carbohydrates into the blood. It appears as if lymphatic adsorption, both of solid materials and of solutions, requires the cooperation of phagocytic cells.

The newer conception of the lymphatics as a closed system is at variance with the older one, in which they were supposed to get smaller and smaller, and their walls less and less complete until ultimately they faded off into the tissue spaces. These, however, bear no closer relationship to lymphatics than they do to blood capillaries. The tissue spaces include all the minute spaces between the fibers and cells of the connective tissues and between the parenchyma of the organs and the great serous cavities of the body (pleural, peritoneal), as well as specially developed tissue spaces, forming the subarachnoid spaces of the brain, the scala vestibuli and tympani of the cochlea and the anterior chamber of the eye. The fluids in these spaces—the tissue fluids—are quite different from the lymph in the lymphatics both in composition and in function. Indeed, the tissue fluids are among the most varied of all the fluids of the body. The spaces may themselves become linked together so as to form a circulatory system, which is quite independent of the lymphatics. This is particularly the case in the brain, where the tissue spaces surrounding every individual nerve cell extend into the subarachnoid area, where they drain into the cerebral sinuses through the arachnoidal villi, which exist as lace-like projections of the arachnoid into the dural sinuses, being covered by a layer of mesothelial cells specially abundant at the tips of the villi, where they form cell nests. Observations of the passage of substances in solution by these pathways have been made by injecting potassium ferrocyanide and citrate of iron into the subarachnoid and subdural spaces and afterwards detecting the presence of the salts by mounting sections in acid media, so as to permit prussian blue to develop. Ordinarily the precipitate is found in or near the villi, but after cerebral anemia it forms in the tissue spaces that surround the nerve cells.

There are therefore three fluids concerned in the transference of food materials and gases between the gastrointestinal apparatus and lungs and the tissue cells—namely, the blood plasma, the tissue fluids, and the lymph. The *tissue fluid,* being in contact with the tissue elements, serves as their immediate nutritive fluid, and it is the function of the blood and lymph to maintain it of proper composition. Everything must be transferred to and from the tissue cells through the tissue fluid, making it

therefore in many ways the most important of the fluids of the body. In the tissue cells themselves there is also the fluid in which the various colloids and crystalloids that enter into the composition of protoplasm are dissolved. This can be removed from cells only by mechanical means, such as grinding with fine sand in a mortar and subjecting the mass to a pressure of several thousand atmospheres in a hydraulic (Buchner) press. This is known as the *tissue juice.* The ultimate exchange of foodstuffs occurs between the tissue fluids and the tissue juices across the cell membrane. The extent and character of this exchange depend on many circumstances, some affecting the cell wall, others, the osmotic and other properties of the two fluids. Obviously, the function of the circulation is to maintain the tissue fluids of correct composition, the blood plasma serving to carry food materials and dissolved oxygen to them (see page 380), but being assisted in the opposite function of removal of effete products by the lymph. The lymph is purely a scavenger; the blood is both purveyor and scavenger.

The above description of the lymphatics is not universally accepted by anatomists, certain of whom believe that the lymphatics are developed from tissue spaces and are consequently much more extensive than they appear to be from injected specimens. The above conclusions are based on reconstruction models, made from serial sections of embryonic tissues, in which the lymphatics frequently appear as isolated vesicles without visible connections. The failure of injections to penetrate into the remoter parts of such a lymphatic system in the embryo is attributed to the discontinuity of spaces, which is, however, removed at later stages of development.

The manner of absorption of injected fluids does not, however, support the view that the lymphatics are directly connected with the tissue spaces. When all the structures of a part are ligated except the main artery or vein, injected poisons which affect central structures, such as the nerve centers, develop their action as quickly as in the intact animal (e.g., strychnine). Similarly, when pigments such as methylene blue are injected into the pleural cavity or subcutaneously, they appear in the urine long before the lymph of the thoracic duct. Such results indicate the pathway of absorption to be the blood rather than the lymph vessels. Through this latter channel absorption proceeds more slowly, but can be greatly assisted by massaging the site of injection. When colored solutions, such as India ink or carmine, are injected subcutaneously, however, a very perfect injection of the neighboring lymphatics may ultimately occur, and through the same pathways microorganisms spread from an infected area.

EXPERIMENTAL INVESTIGATIONS

It has proved a most difficult problem to gain any exact knowledge of the production of lymph by experimental means. Starling, some years ago, in repeating many of the experiments of older physiologists in the light of the newer facts of physical chemistry, added much that is of interest, and it is chiefly with his work that we will concern ourselves here.

The unequal lymph supply of different regions of the body is strikingly demonstrated by comparing the lymph flow from the lymphatics of the leg with that from the thoracic duct. No lymph flows from the former unless the muscles are thrown into activity or the blood is prevented from leaving the limb by ligaturing all the veins. Changes in the arterial blood pressure do not affect the flow. On the other hand, a great increase in the flow from the thoracic duct can readily be induced by disturbances in the blood supply. Obstruction of the portal vein, for example, immediately increases the lymph flow four or five times because of venous congestion in the intestinal capillaries, whilst a still greater increase—perhaps tenfold—is induced by obstruction to the inferior vena cava, which raises the capillary pressure in both the liver and the intestines. After ligation of the hepatic lymphatics (at the hepatic pedicle), obstruction of the vena cava no longer causes the outflow of lymph to increase, indicating that the lymph in the last mentioned experiment must have come from the hepatic lymphatics.

These results, so far as they go, could be satisfactorily explained on the basis that lymph formation is a *filtration* process, that is, a process dependent upon difference in mechanical pressure between the blood capillaries and the tissue spaces. The lymphatics would then serve as channels to return this fluid to the blood vessels through the thoracic duct. The difference in the magnitude of the increased lymph flow from increase in capillary pressure in different regions would be dependent on the permeability of the filter, the capillaries of the limbs being much less permeable than those of the intestine, and particularly of the liver. Another fact in conformity with this view concerns the composition of the lymph from the two regions, that from the limb lymphatics being poor in protein, whereas that from the thoracic duct does not fall far behind the blood plasma in this regard.

Although filtration may explain the considerable increase in lymph flow produced by extreme changes in capillary pressure, it by no means suffices to explain lymph formation under less abnormal conditions. When a muscle or a gland is at rest, it produces practically no lymph,

but during activity the flow becomes marked. This can not be explained by filtration, but may be accounted for by a physico-chemical process—namely, *osmosis*. The energy required for the activity of the tissue cell is produced by chemical changes, whereby large molecules become broken down into numerous smaller ones. These smaller molecules are then discharged into the surrounding tissue fluids, the osmotic pressure of which they increase, with the consequence that water is attracted by osmosis from the plasma in the blood capillaries (see page 4). This increases the volume of tissue fluid, which is then drained away by the lymphatics. The increase in molar concentration will also affect the tissue juices, tending to make the cell swell up by absorbing water. In gland cells this extra water is immediately extruded to form the water of the secretion (see page 421).

An analogous method of lymph formation is not confined to situations where the capillaries are relatively impermeable, for it also occurs in the liver, the lymph flow from which is greatly increased by the injection of bile salts. A similar process no doubt results from muscular activity, although in this case the tissue spaces must form a continuous system of their own, there being, according to most authorities, no lymphatics.

Considerable interest has been taken in the stimulating effect which certain chemical substances have on the secretion of lymph from the thoracic duct. These so-called *lymphagogues* belong to two classes—crystalline and colloidal. Of the former, glucose, urea, and sodium chloride in hypertonic solution, are the best known. Starling explains their action as dependent upon an increase in the osmotic pressure of the blood. This attracts water into the blood from the tissue juices, and leads to an hydremic plethora, with a consequent increase in capillary pressure. If the blood pressure is lowered by hemorrhage before the hypertonic solution is injected, very little stimulation of lymph flow occurs, because there is no available fluid in the tissue to produce the plethora. This observation does not, however, very strongly support the explanation, because so many other disturbances may result from hemorrhage.

The colloidal lymphagogues include watery extracts of the dried tissues of leeches, crayfish, and mussels, as well as commercial peptone. They probably act by damaging the endothelium of the capillaries, so that filtration occurs more readily. Although their action is displayed more particularly on the lymphatics of the liver and intestines, it is also apparent on the skin capillaries, producing cutaneous edema and the formation of blisters (nettle rash).

EDEMA

With such an imperfect knowledge concerning the physiology of lymph formation, it is not surprising that the causes of excessive accumulation of fluid in and between the tissue elements should be little understood. All of the conditions which have been mentioned as capable of causing an increased secretion of lymph—such as increase in capillary pressure, hydremic plethora, action of poisons on the endothelium—are likely to cause edema if the lymphatics of the part are simultaneously obstructed. To produce in animals edema of the subcutaneous tissues like that observed clinically, it is, however, necessary that the vascular disturbance be accompanied either by local damage to the capillary endothelium, such as is produced by arsenic or uranium; or by a general toxemic condition, such as is set up by nephritis. When large amounts of saline solution are injected intravenously, extensive extravasation of fluid may occur into the liver, peritoneum and intestinal lumen, without any subcutaneous edema.

Clinical edemas are of at least three types:

1. The inflammatory edemas, in which the fluid permeates the cells of the inflamed area and does not shift to other parts of the body under the influence of gravity.

2. The nephritic edemas, in which the fluid is more or less loose in the subcutaneous tissues and readily changes its position, and which is accompanied by excess of water in the blood with a corresponding increase of sodium chloride; the percentage concentration of sodium chloride in the blood remains unchanged, but that the other constituents diminished.

3. Cardiac edemas, which are also hypostatic, but are unaccompanied by changes in the relative amount of water and sodium chloride in the blood.

The second and third varieties of edema may of course be more or less present together, for the kidneys are likely to become secondarily affected during venous stasis.

The salt retention in nephritic edema is very significant. As explained elsewhere, it is revealed by comparing the daily output of sodium chloride by the urine with the concentration of this salt in the blood. Less salt is eliminated than would be the case in a normal individual with the same percentage of salt in the blood. In many cases also edema can be diminished by withholding salt from the food. Widal and Javal have conclusively shown the relationship of retention of water in the body, as judged by variations in body weight, to the hydremic condition, as judged by the refractive index of the blood serum, and

to the amount of salt in the diet. A very considerable retention of water usually occurs before there is any evidence of edema; indeed, as a result of giving salt, the body weight may increase from five to seven kilograms (10 to 15 pounds) within a day or two without the appearance of puffiness.

The cause of the edema during salt retention is no doubt closely related to the action of lymphagogues. In a normal person excessive ingestion of salt is immediately followed by excretion of the excess through the kidney. Where the kidneys are diseased, this excess of salt is retained in the blood, raising its osmotic pressure and attracting water from the tissue fluids. This leads to excessive thirst, the imbibed water being used to replace that lost from the tissues. But all the crystalline lymphagogues do not, when present in excess in the blood of nephritic patients, necessarily cause edema; urea, for example, may accumulate considerably without any such effect. The different action is usually attributed to inequality in the diffusibility of the two crystalloids through animal membranes, sodium chloride diffusing much less readily than urea.

It is most important to note that the fluid in edema is loose in the tissues and can be drained away by the insertion of tubes. There is absolutely no evidence in support of the claim of Martin Fischer that edema is due to imbibition of water by the colloids of the tissues. This question has been fully discussed elsewhere (page 62).

BLOOD AND LYMPH REFERENCES

(Monographs)

[1]Howell, W. H.: The Harvey Lectures, J. B. Lippincott Co., xii, 272.
[2]Starling, E. H.: Human Physiology, Lea & Febiger, 1915.
[3]Rowe, A. H.: Arch. Int. Med., 1917, xix, 354.
[4]Williamson, C. S.: Arch. Int. Med., 1916, xviii, 505.
[5]Tower and Herm: Proc. Soc. Biol. and Med., 1916, xviii, 505.
[6]Rous and Robertson: Jour. Exp. Med., 1916, xxiii, 219, 239, 549.
[7]Butler, G. G.: Quart. Jour. Med., 1912, vi, 145.
[8]Howell, W. H.: cf. Harvey Lecture; also Am. Jour. Physiol., 1913, xxxii, 264.
[9]Drinker, C. K., and K. R.: Am. Jour. Physiol., 1916, xli, 5.
[10]Denny and Minot: Arch. Int. Med., 1916, xvii, 101; Am. Jour. Physiol., 1915, xxxviii, 233.
[11]Addis, T.: Quart. Jour. Med., 1910, iv, 14.
[12]Cannon and Mendenhall: Am. Jour. Physiol., 1914, xxxiv, 225.
[13]Howell, W. H.: Arch. Int. Med., 1914, xiii, 80.
[14]Brodie, T. G.: Jour. Physiol., 1897, xxi, 403.
[15]Whipple, G. H.: Arch. Int. Med., 1912, ix, 365; Jour. Exp. Med., 1911, xiii, 136.
[16]Whipple, G. H.: Arch. Int. Med., 1913, xii, 637.
[17]Duke, W. W.: Arch. Int. Med., 1912, ix, 258.

PART III

THE CIRCULATION OF THE BLOOD

CHAPTER XV

BLOOD PRESSURE

The object of the circulation is to maintain through the tissues a supply of blood that is adequate to meet their demands for nutriment and oxygen and to remove the effete products of their metabolism. The demands vary according to the activities of the tissue, being particularly variable in the case of such tissues as the muscular and the glandular. In studying the physiology of the circulation we have therefore to bear in mind two aspects of the problem: (1) the cause for the continuous bloodflow, and (2) the mechanism by which alterations in this bloodflow are brought about.

If we open an artery we shall find that the blood escapes from it under such a pressure that it is thrown to a height of about six feet, that its outflow is proportional to the size of the artery, and that it pulsates. If, on the other hand, we open a vein, we shall find that the blood wells out without any very evident pressure, and that it flows in a continuous stream, its outflow being the same in a unit of time as that of the artery, provided the two vessels are the only ones supplying the particular area. The general conditions governing the bloodflow are the same as those governing the flow of fluid through any system of tubes. For example, in the city water mains it is known to every one that the rate of outflow from any part of the system depends finally on two factors: (1) the difference in pressure at the beginning and end of the system, and (2) the caliber of the tube at the outlet. We may increase the outflow by raising the pressure at the beginning of the system, the caliber of the outlet meanwhile remaining constant, or we may maintain the pressure constant but increase the caliber of the outlet.

In the circulation of the blood, the difference in pressure at the beginning and end of the circulation is furnished by the pumping action of the heart, and the alteration of the caliber of the outlet is provided for by the constriction or dilatation of the blood vessels. These simple physical principles indicate the direction which a study of the circulation

should take. They indicate that our first consideration should be of the mean blood pressure, how it is maintained, and how it can be made to vary. After we have learned this, we may then proceed to a more particular examination of the mechanism of the pump—that is, of the heartbeat; then finally we may proceed to examine the nature of the processes by which the caliber of the arteries is controlled.

THE MEAN ARTERIAL BLOOD PRESSURE

The first prerequisite to the investigation of the blood pressure, as of any other physical problem, is that we should possess some means by which it can be quantitatively measured. The earliest attempt to accomplish this was made by the English scientist, the Rev. Stephen Hales, a little over a century after Harvey published his account of the circulation of the blood. Hales connected a glass tube nine feet in length with a severed artery of a horse, the connection between the two being made by means of a piece of brass pipe joined to the windpipe of a goose as a substitute for rubber tubing. He found on untying the ligature on the artery that the blood rose in the tube to a height of eight feet and three inches above the level of the left ventricle of the heart, and that when at full height it rose and fell with each pulse through a distance of two, three or four inches.

Mercury Manometer Tracings

The somewhat crude but very significant experiment of Hales clearly established the existence of the enormous pressure at which the blood is made to circulate through the arteries. To render possible a further investigation of the factors on which this pressure depends, it became necessary to invent some more convenient means for its measurement, but this was not accomplished until a century later, when Poiseuille applied the *mercury manometer,* which Ludwig subsequently adapted so that tracings might be taken (Fig. 21).

Having before us such a tracing as shown in Fig. 22, let us consider how it may be used in the study of blood pressure. The first thing we must do is to measure the average height of the tracing above the line of zero pressure; the mean arterial blood pressure is then equal to this distance multiplied by two, because the distance through which the mercury has moved up in the limb of the manometer carrying the writing point is only one-half of its total displacement. Since mercury is about 13.5 times heavier than an equal volume of blood, the above measurement must be multiplied by this figure if we desire to express

our result in terms of the height to which the blood pressure could raise a column of blood.

In arteries of approximately the same size, the mean arterial blood pressure does not markedly vary in different mammals. Thus, in the carotid artery of the dog it averages about 110 to 120 mm. Hg, in that of the cat about 105 to 115 mm., in the rabbit from 90 to 105 mm., in the sheep about 150 mm., in the horse about 200 mm., and in man some-

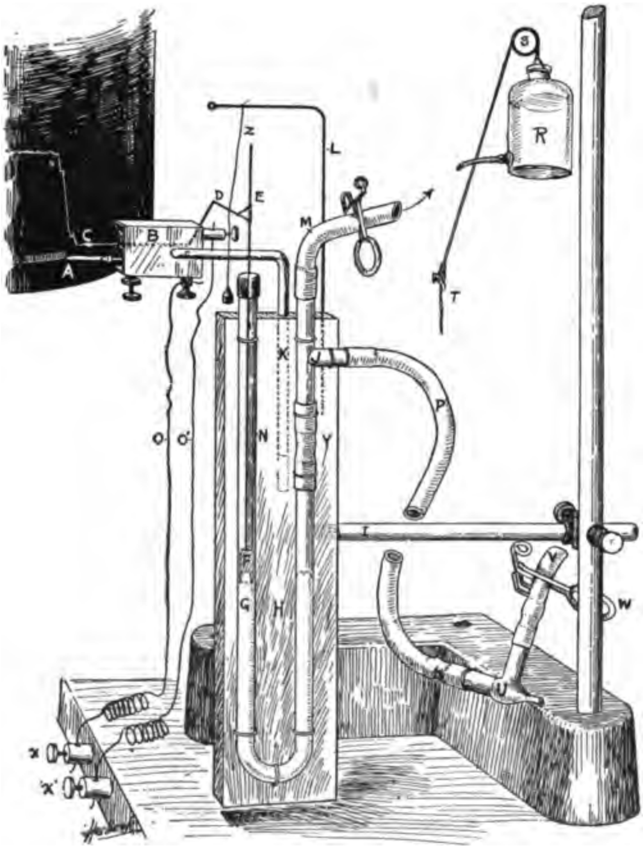


Fig. 21.—Mercury manometer and signal magnet, arranged for recording the mean arterial blood pressure in a laboratory experiment. The pressure bottle (R) is filled with anticoagulating fluid and is connected by tubing with the manometer (M), the cannula for the artery (U) being connected with the T-piece (J). By this arrangement it is possible to flush out the tubing when clotting interferes with the experiment. (From Jackson—Experimental Pharmacology.)

where between 120 and 140 mm. The pressure varies in different parts of the vascular system, being greatest in the aorta and least in the smallest arterioles but the fall in pressure—the pressure gradient—does not become very pronounced until the arterioles have become so small that it is no longer possible to insert a cannula into them; thus, the mean

blood pressure in the renal or femoral artery is very little less than that in the aorta.

If we examine *the contour of the tracing* which the pressure draws, we shall find that it exhibits two types of wave, small and large; and if we observe the animal while the tracing is being taken, we shall find that

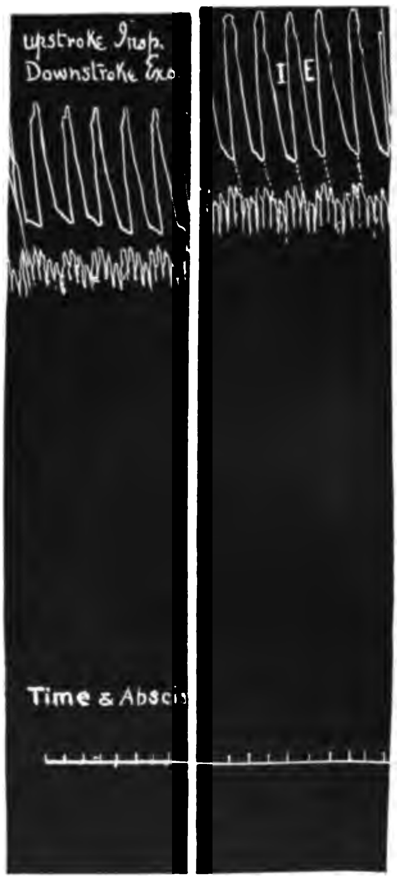


Fig. 22.—The arterial blood pressure recorded with a mercury manometer (lower tracing), along with a tracing of the respiratory movement of the thorax. Note that the beginning of respiration occurs distinctly before the rise in blood pressure.

the former are caused by the heartbeats and the latter by the respirations—an observation which immediately raises the question as to the trustworthiness of the method, for it will be asked, How can it be that the heartbeat produces an effect on blood pressure which is less than that of the respirations? Obviously the tracing must be faulty in regard to the relative significance of the waves.

Spring Manometer Tracings

The cause of this inaccuracy depends on the inertia of the mercury, an inertia which is so great that the sudden changes of pressure produced by each heartbeat are not able to overcome it, whereas the much less significant but more prolonged pressure changes produced by each respiration develop their full effect on the mercury. These facts led investigators to seek for instruments in which the inertia error is eliminated, with the result that they invented what are known as *spring manometers.*

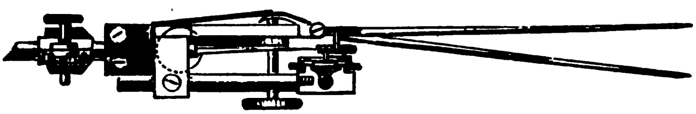


Fig. 23.—Hürthle's spring manometer.

Many varieties of this instrument have been produced, but for our purpose it is necessary to describe the principle of only the simplest and most efficient—the Hürthle manometer. As shown in Fig. 23, it consists of a variety of tambour, which differs from the ordinary tambour in two important particulars: (1) its chamber is made as small as possible, and (2) it is covered not with an elastic membrane but with one of leather or of thin fluted metal. These two precautions are taken in order to avoid spurious waves set up on account of elastic recoil. Such errors are further reduced by filling the tubing and chamber of the tambour with an anticoagulating fluid.

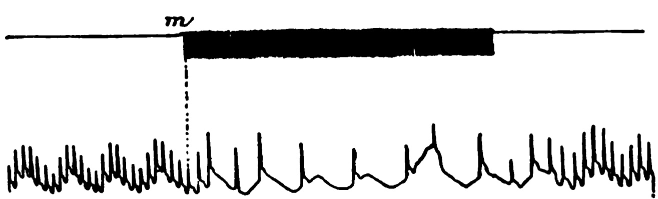


Fig. 24.—Arterial pressure recorded by a spring manometer. The effect of weak excitation of the vagus is seen during the period marked by the signal *m*. (From Dubois.)

Before the tracing taken with the spring manometer can be employed for quantitative measurements, it must obviously be graduated according to some scale. This is accomplished immediately before or after the experiment by connecting the manometer through a T-piece with a pressure bottle, which can be raised or lowered to a specified height, and with a mercury manometer. The displacement of the writing point of the spring manometer corresponding to each 10 mm. Hg of pressure is then written on the tracing.

The tracings taken with such a manometer, as shown in Fig. 24, are quite different from those with the mercury manometer. It will be seen that now the cardiac waves are decidedly the more pronounced, the respiratory, being comparatively inconspicuous. Instead of there being a fairly steady pressure in the arteries, this undergoes very considerable alteration during each heartbeat.*

Examination of this tracing gives us accurate information regarding the blood pressure both between the heartbeats—diastolic, as it is called—and during them—systolic. It gives us a means of telling what must be the *dead* load of the circulation—that is, the pressure that is constantly present—as well as the *live* load that is superadded to this by each heart-

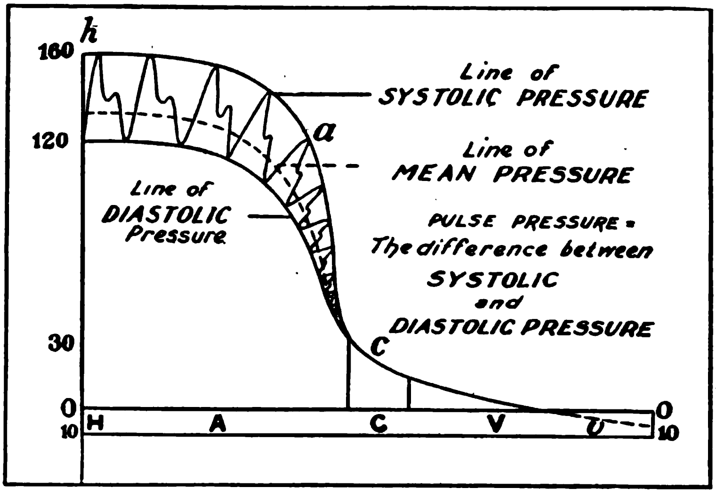


Fig. 25.—Diagram based on experiments on dogs to show the magnitude of the systolic, diastolic and mean blood pressures at different parts of the circulatory system. O is the line of zero pressure, and the letters below it indicate the parts of the system to which the curves refer. (From Brubaker.)

beat. This difference is often called *the pressure pulse,* and in man it amounts to somewhere about 35 mm. Hg. If we take tracings with a spring manometer from different parts of the arterial tree, we shall find that, as we travel towards the periphery, the pressure pulse becomes less and less marked, until finally by the time the capillaries are reached it has almost entirely disappeared. This decline in the pressure pulse can moreover be seen to be dependent more largely on a fall in systolic than in diastolic pressure. In other words, the dead load of the circulation—the diastolic pressure—remains practically constant all along the arterial tree, whereas the systolic pressure falls relatively quickly (Fig. 25).

*The tracings shown in Figs. 22 and 24 are not typical, the pulse pressure being too small in the latter and too large in the former.

Clinical Measurements

The methods of blood-pressure measurement in man have recently become so perfected that the results are almost as accurate as those obtained in laboratory animals by direct measurement through the use of cannulæ inserted into the vessels. Both the systolic and the diastolic pressure can be measured with equal facility and accuracy. Since the technic for making the systolic measurements was described at a much earlier date than that for the diastolic, it has until recently been the habit with a great part of the medical profession to be satisfied with systolic readings alone. This is most unfortunate, because the knowledge which such information gives us is incomparably inferior to that which can be obtained by gauging the diastolic pressure. Until we have learned more about the dynamics of circulation, it would be profitless to go into any details as to the reasons for this statement, but it will soon become self-evident. Suffice it for the present to state that *the diastolic pressure is the more important because it gives us the load which the vessels and aortic valves must constantly bear,* and the resistance which must be overcome prior to the opening of these valves at the beginning of systole. Moreover, it helps us to gauge the peripheral resistance.

The first step in the technic of blood-pressure measurements in man is the placing of an armlet or cuff around the arm or leg. This armlet consists of a rubber bag at least 12 cm. broad and covered on its outer surface by cloth or leather. The bag is connected by tubing with a pressure gauge and a pump. The pressure gauge may be either an ordinary mercury manometer or one of the numerous gauges built on the aneroid principle that are now on the market (Fig. 26). For measuring the blood pressure in the vessels of the upper extremities, the armlet should be applied around the fleshy part of the upper arm and for the lower limbs around the thigh. For accurate reading of both pressures the following *procedure* should be followed. Having applied the armlet, the pulse should be palpated at the radial artery, and the pressure in the armlet then raised until the pulse can no longer be felt, at which moment the pressure in the manometer should be noted. The cuff should then be slowly decompressed and the pressure noted at which the pulse reappears. These two readings of systolic pressure should be close together, but they will not usually agree exactly for reasons which will be explained immediately. They give us *the palpatory systolic index,* as it is called. The pressure is now lowered about 15 mm. Hg, and a stethoscope is placed in front of the bend of the elbow over the artery and as close up to the cuff as possible. With each heartbeat a distinct sound like a pistol shot will be heard. The decompression is now continued slowly, and as the pressure falls the sounds will be heard to become louder and prob-

ably somewhat murmurish in quality. At a certain pressure this loud character of the sound will suddenly become much less marked, and the murmurish quality if present will suddenly disappear. This point corresponds to the diastolic pressure, which is now read off from the manometer.

It must be remembered that below this point, as the pressure in the cuff is further lowered, a sound is still heard in the artery; indeed it does not entirely disappear until the pressure has become quite low. This point of final disappearance is, however, of no significance. The cuff is

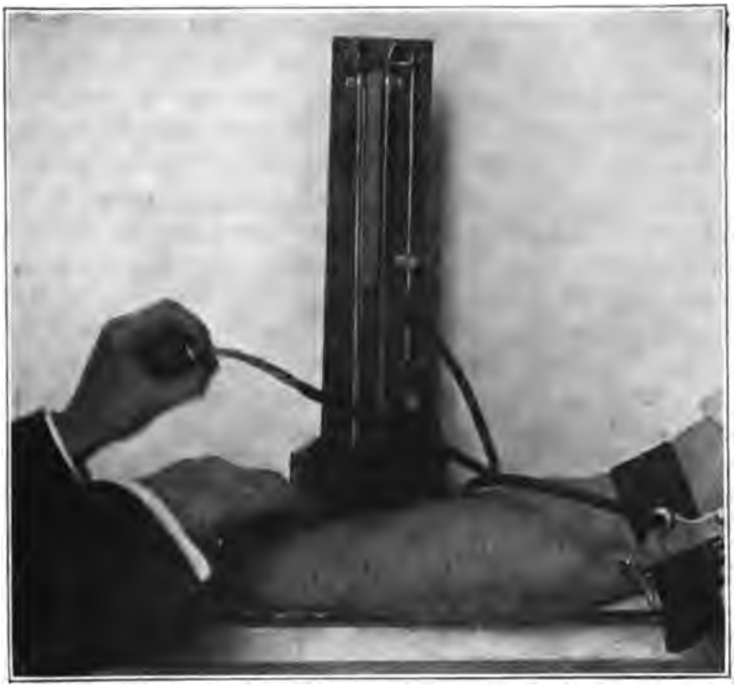


Fig. 26.—Apparatus for measuring the arterial blood pressure in man. The pressure in the cuff is raised by means of the syringe until the pulse can no longer be felt at the wrist. This pressure is read off on the mercury manometer (systolic pressure).

now entirely decompressed, and should be left so for a moment or more, so that the circulation in the part of the arm below it may return to the normal.

The above readings should then be controlled by a second observation, in which the methods employed are slightly modified. With the stethoscope at the bend of the elbow the pressure in the cuff is run up to a little above the previously determined diastolic pressure, so that the sound is clearly heard. The pressure is then further raised till the sound disappears. This point indicates the systolic pressure; it is called

the *auditory systolic index*. It will be found to give a systolic pressure a little *higher* than that obtained by palpation of the artery at the wrist. The sound being now absent, the pressure in the cuff is lowered until the sound reappears, and the point at which this occurs should almost exactly correspond to that at which the sound was found to disappear. If the palpatory systolic index is not below the auditory, it indicates that some error has been made in the application of the apparatus, and that the reading of the diastolic pressure will be unreliable. The usual source of error is in the position of the stethoscope. If readjustment of this does not bring the two indices into proper relationship, the auscultatory method can not be relied upon for either systolic or diastolic readings.

In case of the failure of the auscultatory method, we have to fall back upon the palpatory method for systolic pressure; and for the measurement of diastolic, we have to use the method known as the *oscillatory*, which until recent years was the only one used for gauging the diastolic pressure. This consists in observing the oscillation of the indicator of the pressure gauge; as the pressure in the cuff falls gradually from below the systolic pressure, these oscillations will be observed to increase in amplitude, until they reach a maximum beyond which with lower pressure they rapidly decline. The pressure in the cuff at the moment when the oscillations are at the maximum represents the diastolic pressure. With a mercury instrument it is obviously difficult to employ this method, but with a modern spring instrument it can with a little practice be used with great accuracy and will serve as a valuable check on the diastolic reading as taken by the auscultatory method.

The procedure may be altered in various ways, there being only one precaution to bear in mind; namely, that the pressure in the cuff should not be applied continuously for more than a few moments of time, for if this is done for long periods, not only will it interfere with the accuracy of the reading, but it may cause considerable discomfort to the patient.

There are several conditions affecting *the accuracy of the readings* by each method which it is well to bear in mind. These have been investigated by MacWilliam,[1] Leonard Hill,[2] and Erlanger.[3] With regard to *the systolic pressure* the most important of these are as follows: (1) The compression cuff should be a wide one (12 cm.), and it should never be applied so that there is any chance of its compressing the artery against a bony surface. This precaution is necessary, since it has been found that much less pressure is required to obliterate any perceptible pulse below the armlet when the artery is flattened against some hard structure than when it is uniformly compressed in the tissues in which it lies. (2) Discrepancies are often noted between the systolic readings on compres-

sion and decompression of the artery; that is, the pulse may reappear on decompression at a lower pressure than that at which it disappeared on compression, the difference being most marked when the decompression is done quickly. This difference is owing to the fact that the full force of the pulse does not reach the forearm until all the vessels have become distended with blood. (3) There are often discrepancies in the systolic readings taken from different limbs; thus, it is not uncommon to find that the systolic pressure in the leg is higher than that in the arm even when the observed person is in the horizontal position. These differences are most commonly observed in patients suffering from aortic regurgitation or thickened arteries. In aortic regurgitation the pulse is of the water-hammer variety, and the greater systolic pressure observed in the leg vessels in such cases seems to depend on differences in the physical conditions concerned in the transmission of this exaggerated pulse wave to the vessels of the two extremities.

The reason for the discrepancies in cases of hardened arteries is no doubt that the hardening is likely to be more pronounced in the vessels of the thigh than in those of the arms. When a hardened vessel is compressed it does not collapse uniformly—that is, it does not become completely closed—but its walls come together at the middle part while chinks still remain at the sides. The blood continues to pass through these chinks, and a very considerably higher pressure in the cuff is required to obliterate them. That this is probably the correct explanation is supported by the observation that, although in such patients the pulse does not disappear in the vessels of the foot at the same pressure as it does at the wrist, a distinct change is nevertheless perceptible in the pulse of the foot at a cuff pressure equal to that producing obliteration in the wrist. In a patient showing a systolic pressure of 115 mm. for the upper arm and 198 mm. for the leg, at 116 mm. the pulse in the leg, although not obliterated, became notably cut down in volume. It thereafter persisted at a small volume with little alteration until the pressure became sufficient to obliterate it. It is said that repeated compression and decompression of the hardened arteries greatly reduces the discrepancy in the systolic readings. Differences in systolic readings are also sometimes observed in normal individuals, particularly after muscular exercise, but for these no satisfactory explanation can be given.

While palpating the radial artery, it will often be noticed, as the pressure in the cuff is gradually raised from zero, that the force of the pulse increases perceptibly until a pressure of about 50 mm. is reached. This paradoxical behavior of the pulse can also be demonstrated by the sphygmograph (see page 201). Its cause is not understood, but it is of significance that the greatest augmentations occur at a cuff pressure

at which a sound first comes to be heard by listening over the artery at the elbow.

With regard to the *diastolic pressure,* there has been some controversy as to whether it is more accurately gauged by the oscillatory or the auscultatory method. If both methods are employed it will usually be found that the oscillatory gives a higher reading than the auscultatory. The concensus of opinion seems to be that the latter method is the more accurate, and certainly it is the easier to apply, for with the oscillatory there is often great difficulty in deciding just exactly when the maximum oscillation occurs.

The strongest evidence supporting the conclusion that the auscultatory readings are more reliable than the oscillatory has been gained by experiments with an artificial schema, consisting of a wide glass tube representing the armlet, filled with Ringer's solution and closed by rubber stoppers pierced by tubes, which are connected with a fresh artery, which therefore runs from end to end inside the tube. Through tubing connected with the artery a pulsatile flow of oxygenated Ringer's solution is made to flow at varying pressures, which are indicated by valved manometers (see page 152) connected with the artery tubing just beyond the compression tube. The pressure in the latter is also measured by a manometer, and it is caused to vary by a suitable compressor. By comparing the behavior of the artery with the pulsating movement of a spring manometer connected with the compression chamber, under different degrees of pressure inside and outside the artery, it has been observed that the maximal oscillation occurs when the artery is actually somewhat flattened between the pulse beats; that is, it occurs at an outside pressure above the diastolic pressure, at which of course the vessel should retain its circular shape. When a stethoscope is applied to the tube leading from the artery just beyond the compression chamber, in the above described model sounds similar to those in the arm are heard with each pulsation. While the pressure is being gradually lowered from above the obliteration point, these sounds will be found to become first audible as soon as a certain amount of fluid is forced through the compressed area at each pulse (the systolic index), and to become louder and often murmurish in quality as the decompression is proceeded with, until a pressure is reached at which they suddenly become less intense and change in character. At this moment it will be observed by watching the artery that the external pressure is no longer capable of producing any flattening of the vessel between pulses. Evidently, therefore, the change of sound corresponds exactly to the diastolic pressure.

With regard to the cause of the sound, it should be clearly understood that it is the systolic wave that produces it, but its occurrence and its

character when present are dependent upon the intra-arterial pressure existing during the diastolic phase. The cause of the sound has been shown to depend on the production of a water-hammer in the blood vessels below the compression cuff (Erlanger[3]). By a water-hammer is meant the pressure changes which are caused by suddenly stopping the flow of water in a pipe. When a sudden pressure occurs in vessels with elastic walls, these walls are thrown into vibration and so produce a sound. In the taking of blood-pressure measurements, as above described, when the pressure in the cuff is between systolic and diastolic, the volume of the compressed artery will increase abruptly with each heartbeat and thus permit a considerable volume of swift-flowing blood to enter the rest of the artery *underneath* the cuff. When this quickly moving column of blood comes into contact with the stationary blood filling the uncompressed artery *below* the cuff, it will become immediately checked, and thus distend the arterial wall with unusual violence and set it into vibration.

CHAPTER XVI

THE FACTORS CONCERNED IN MAINTAINING THE BLOOD PRESSURE

Having become familiar with the principles of the methods by which blood-pressure measurements are made, the next problem is to examine into the causes which operate to maintain the pressure. Two of these causes may be considered as fundamental, since without them no such pressure could exist. These are: (1) the pumping action of the heart, and (2) the peripheral resistance—that is, the resistance to outflow of blood from the ends of the arterial system. Less essential though important factors are: (3) the volume of blood in the blood vessels, (4) the viscidity or viscosity of the blood, and (5) the elasticity of the walls of the vessels. We shall now proceed to examine the experimental evidence which indicates the relative importance of each of these factors.

1. The Pumping Action of the Heart

Changes produced in the mean arterial blood pressure by alteration in the pumping action of the heart are most strikingly demonstrated by observing this pressure after cutting or during stimulation of the vagus nerves. As will be explained later (page 217), impulses conveyed through these nerves to the heart make the beats slower and weaker. These impulses are constantly acting in the heart, so that when both vagus nerves are cut, the beats become more frequent and stronger, with the result that the mean arterial pressure rises considerably. A lesser degree of this effect can usually be obtained by cutting the vagus nerve on one side (Fig. 27). If now the peripheral end of a cut vagus nerve is stimulated, as by applying an electric current to it, the heart will either stop beating altogether or become very much slowed, with the result that the mean arterial blood pressure will fall, in the former case almost to zero and in the latter, to a level corresponding to the degree of slowing of the heart (Fig. 28).

2. The Peripheral Resistance

To demonstrate the influence of peripheral resistance on mean arterial blood pressure, the most striking experiment is performed by cutting or stimulating the great splanchnic nerve. In this nerve impulses,

which are called vasoconstrictor because they constrict the lumen of the blood vessels, are transmitted to the blood vessels in the abdomen. The vessels are under the constant influence of these impulses so that, when the nerves that transmit them are severed, the vessels dilate and thus offer less resistance to the movement of blood along them. The result produced on the mean arterial blood pressure by cutting the two splanchnic nerves is therefore a marked and sudden fall, which is immediately recovered from if the peripheral end of one of the cut nerves is stimulated artificially (Fig. 29). In choosing this experiment to prove the relationship between peripheral resistance and the mean arterial blood

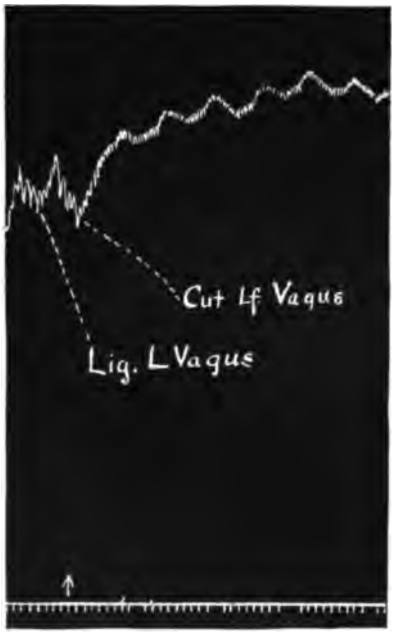


Fig. 27.—Effect of cutting the vagus nerve on the arterial blood pressure.

pressure, it must be remembered that it is not entirely conclusive, since the results observed on the mean arterial blood pressure from cutting or stimulating the nerve may be in part explained as due to variation in the total capacity of the circulation; more room is created by cutting the nerves, less room by stimulating them.

3. The Amount of Blood in the Body

This can be altered by hemorrhage or transfusion, and the results of such procedures are of interest not only on account of their physiologic bearing, but also because of their great practical importance.

To appreciate the significance of the results, it is important to bear in mind that *the total volume of the blood* constitutes from 5 to 7 per cent of the weight of the animal. This fact has been determined partly by postmortem, and partly by antemortem measurements. In the postmortem method, the total amount of blood is determined by collecting the blood while bleeding the animal to death and then washing out the vessels with saline solution until the escaping fluid is no longer tinged with red. The blood contained in the saline solution is estimated by colorimetric methods (see page 92), and is added to that directly col-

Fig. 28.—Effect of stimulating the peripheral end of the right vagus on the arterial blood pressure.

lected. In the antemortem method some substance that does not diffuse through vessel walls or become quickly destroyed is added to the blood. By determining the concentration of this substance in a specimen of blood, the volume with which it has become mixed can readily be calculated. Acacia has recently been found suitable for this purpose (Meek), but the best known work (of Haldane) was done by causing the animal to inspire a known amount of carbon monoxide. This combines with the hemoglobin of the blood (see page 401) to displace an equal quantity of oxygen. By determining the difference between the volume

of carbon monoxide in the blood before and following its administration we can calculate with how much blood the known inspired quantity of carbon monoxide must have combined. The results vary somewhat in different animals; in the dog, the blood constitutes about 7.7 per cent of the body weight, and in man, about 5 per cent.

The immediate effect of hemorrhage on the blood pressure depends on the rate of bleeding. If a large artery, such as the femoral, is cut across,

Fig. 29.—Effect of stimulation of the left splanchnic nerve on the arterial blood pressure. Note the primary and secondary rises.

the pressure will show an immediate but moderate fall, due largely to the fact that we have suddenly decreased the peripheral resistance. If on the other hand only a small artery or a vein is opened, the bleeding will at first produce no effect on the blood pressure, and it is only after some considerable amount of blood has been removed that it begins to fall (Fig. 30). To be more exact, we may state that the removal of 5 c.c. of blood per kilogram of body weight does not influence the blood pressure. The removal of a second portion of 5 c.c. per kilogram causes the blood pressure to begin to fall, the fall of pressure for each subsequent 5 c.c. of

blood per kilogram removed averaging about 6 mm. Hg, until after 20 to 25 c.c. of blood per kilogram have been removed, when a more rapid fall in pressure sets in (Downs[4]). When the pressure reaches the level of from 20 to 30 mm. Hg, the danger limit is reached, for there now supervenes a train of symptoms known as "shock," and the chances for the animal's recovery become uncertain. That the removal of the first portion of blood, if this removal is slow enough, does not influence the blood pressure, indicates that some adjustment has occurred in the vascular system to hold up the pressure in spite of the loss of blood. This adjustment is believed to consist in vasoconstriction.

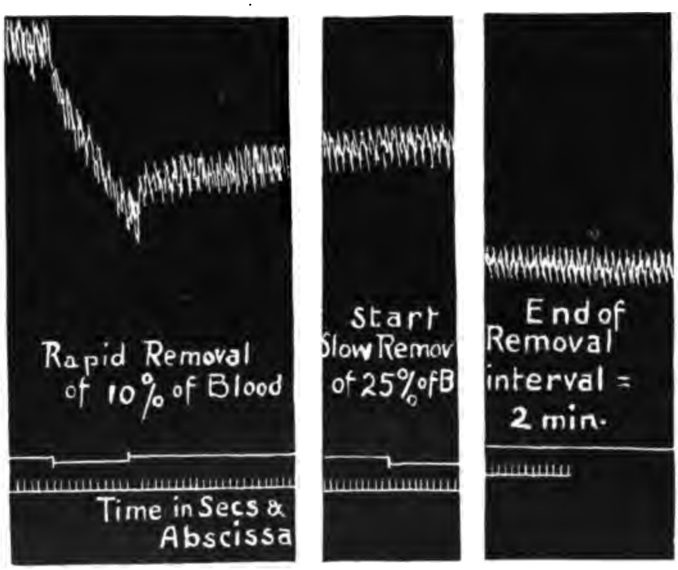


Fig. 30.—The effect of rapid and slow hemorrhage on the arterial blood pressure. Between the second and third pieces of tracing an interval of two minutes elapsed.

Recovery from hemorrhage is remarkably rapid, the original volume of blood being restored within a few hours. The chances of recovery depend upon the amount of blood lost. A loss equal to 2 or 3 per cent of the body weight can almost always be recovered from in laboratory animals, and in the case of man there is reason to believe that recovery may occur after as much as 3 per cent of the body weight has been lost. The recovery of blood pressure is brought about partly by a transfer of fluid from the tissues to the blood. This abstraction causes a drying out of the tissues, which soon excites an extreme degree of thirst. The dilution of blood by fluid derived from the tissues occurs very rapidly, as can be shown by comparison of the hemoglobin content, or the number of blood corpuscles, in samples of blood removed immediately before

and immediately after a hemorrhage. The specific gravity of the post-hemorrhagic blood is also decidedly below normal, indicating that the diluting fluid contains a lower concentration of dissolved substances than the blood plasma. The dilution of the blood is indeed often so great that hemolysis occurs, the plasma being distinctly tinted red.

Hemorrhage also slightly raises the hydrogen-ion concentration of the blood plasma, and diminishes the store of reserve alkali, so that the addition of a certain amount of acid to the blood (e.g., carbon dioxide) causes a greater rise in the hydrogen-ion concentration.

The deficiency in the blood elements produced by the dilution is rectified by the manufacture of new corpuscles in the bone marrow, etc., but this process in a liberally fed animal takes several days for accomplishment, and while it is going on microscopic examination of the blood will reveal the presence of immature corpuscles.

Careful studies of blood regeneration following the removal on two successive days, of 25 per cent of the blood, by Whipple and Hooper, have shown that even in starving animals the total amount of hemoglobin (percentage of hemoglobin multiplied by the volume of blood) slowly recovers. Recovery is greatly hastened by feeding with flesh or even with gelatin. Removal of the spleen or the establishment of a biliary fistula does not interfere with the recovery.

Incidentally it will be advantageous to consider here the **effects of transfusion.** These are very different according to the nature of the fluid used for transfusion. Three transfusion fluids have been investigated: (1) blood itself, (2) physiologic saline solution (see page 95), and (3) physiologic saline solution containing viscid substances such as gelatin. The effects are also very different according to whether the solutions are injected into animals with normal blood pressure or into those whose blood pressure has been lowered by preceding hemorrhage.

When blood is injected into animals with normal blood pressure, it will very soon cause the pressure to rise, and as the injection is maintained the rise may continue until the pressure is perhaps 50 per cent or more above its normal level. If the injection is long continued, however, a sudden fall of pressure occurs, on account of engorgement of the right side of the heart. If the injection is not pushed so far, the increased blood pressure after being maintained for a short time returns to its old level.

Injection of saline into a normal animal, if made slowly, has no effect at all on the blood pressure; if more rapidly injected, the pressure will rise slightly, but to a much less extent than that observed when blood itself is injected. Much larger quantities of the saline than of the blood can be tolerated before cardiac embarrassment ensues. After the dis-

continuance of the saline injection, the blood pressure returns very rapidly to its old level. The most striking result of such experiments is the enormous volume of saline solution which can be slowly injected without perceptibly affecting the pressure. The question is, Where does the fluid go? If the urinary outflow is examined, a certain increase will usually be observed, but never by any means sufficient to account for the disappearance of the injected saline. If we open the abdominal cavity, we shall find that a considerable transudation of the saline into the peritoneal cavity has occurred, and that the liver is conspicuously edematous. A certain degree of edema is also usually evident in the tissues of the extremities.

Still more interesting and important, from a practical standpoint, are the results obtained by injecting the above solutions into the animals whose blood pressure has been lowered by a previous hemorrhage. If the blood removed during the hemorrhage is defibrinated (see page 101), and then reinjected into the animal, it will bring the blood pressure almost but not quite back to its original level, which will then be fairly well maintained. If, on the other hand, saline solution instead of blood is injected, the restoration of blood pressure (with an amount of saline equal to that of the removed blood) will amount to only about three-quarters of the extent to which it had fallen. This partial recovery is, moreover, maintained for only a short time, after which the pressure rapidly falls nearly to the level to which it was reduced by the hemorrhage.

These observations raise two important practical questions: (1) Why is saline relatively ineffective in the restoration of pressure? and (2) Why does the pressure thus restored so quickly fall again?

The answers to these questions brings us to a consideration of the next of the factors concerned in the maintenance of the blood pressure, namely, the viscosity of the blood.

4. The Viscosity of the Blood

The importance of this factor arises from the fact that facility of flow in a tube is inversely proportional to the viscosity of the fluid and directly proportional to the driving pressure to which it is subjected—that is, to the difference in pressure between two points in the tube. If therefore the output of the heart should remain constant, but the viscosity of the blood be decreased by a saline injection, the facility of flow would be increased and the pressure decreased. This fact can easily be shown experimentally in a model by causing gum solutions of various concentrations to be driven through a glass tube by means of a small piston pump delivering a constant amount of fluid into the tube with

each movement. Although the outflow from the narrow end of the tube must remain constant, the pressure in the tubing will vary in proportion to the viscosity of the gum solution (Bayliss[5].)

Transferring these results to an animal whose blood pressure has been lowered by hemorrhage, it has been found that if saline solutions containing a sufficient amount of gum acacia or gelatin to make the viscosity about equal to that of blood, are injected, the original level of blood pressure is recovered as well as it would be had blood itself been injected. A 7 per cent solution of gum acacia almost fulfills these requirements, but unfortunately this solution contains a slightly greater amount of calcium than it is safe to inject into an animal. The excess of calcium may, however, be removed by exactly neutralizing the gum solution with sodium hydroxide, neutral red being used as an indicator. Most of the calcium becomes precipitated as phosphate. The mucilage of the British Pharmacopeia, diluted five times with water, makes a 7 per cent solution of gum acacia. A 6 per cent solution of gelatin, after being heated to 100° C., gives a viscosity similar to that of blood, but on account of the possible presence of tetanus spores such solutions must be very carefully sterilized before injection, and the process of sterilization causes a decrease in viscosity. The injection of a quantity of one of the above solutions equal to that of blood lost by a hemorrhage will usually bring the blood pressure back to its original height and hold it there for an hour or so.

Viscosity is, however, not the only property of such solutions upon which their desirable effect depends. The *osmotic pressure of the colloids* also comes into play. By injecting saline solution containing a sufficient amount of a colloid such as soluble starch, which gives it the correct viscosity but has no osmotic pressure, the blood pressure, although it temporarily recovers after transfusion, does not maintain its recovery in the same way as with solutions containing gum or gelatin. The difference between a starch solution and one of gum or gelatin is that the former has no osmotic pressure, the effect of which is developed mainly on the excretion of urine, as can be shown by observing the outflow from the ureters during the injection into animals of equal quantities of saline alone or of saline containing starch or gelatin (Knowlton[6].) With the first two fluids diuresis is produced, but not with gelatinous solutions. The reason that the osmotic pressure of certain colloids prevents passage of water from the blood into the uriniferous tubules is that the development of this pressure on the blood side of the renal epithelium tends to counteract the filtration pressure by which the urine is formed (see page 514.)

Although the urinary factor will not in itself explain the efficiency of

the colloids in recovering the blood pressure, the conditions controlling it reveal the mechanism by which the passage of fluid from the blood vessels into the tissues is prevented when solutions of correct composition are injected. Normally the protein content of the blood plasma is higher than that of the tissue lymph, so that there is a continual attraction of water from the tissues to the blood—an attraction which is normally balanced by filtration going in the opposite direction. When the filtration pressure in the blood vessels exceeds the difference existing between the osmotic pressure of their contents and that of the tissue fluids, water will pass into the tissue spaces. When the blood is diluted, as by the injection of saline solution, the osmotic pressure of the colloids in a given volume becomes lowered and, the filtration pressure remaining constant, fluid passes into the tissue spaces. Of course these explanations rest on the assumption that the wall of the blood vessels consists of a membrane which is permeable to crystalloids but impermeable or nearly so to colloids.

Another important property of the transfused saline solution to consider is its *hydrogen-ion concentration.* This value increases in the blood left in the body after hemorrhage, and injection of sodium chloride solution aggravates the acidosis; addition of $NaHCO_3$ so as to make a 0.2 M solution restores the correct P_H, and at the same time restores the lost buffer influence (Milroy[7].) These observations are of interest in the light of the recent discovery of Cannon that a condition of acidosis, as judged by the CO_2-combining power of the blood, is present in shock, and that the development of this condition can often be guarded against by bicarbonate injections.

5. Elasticity of Vessel Walls

The elasticity of the vessel walls is essential to the maintenance of the diastolic pressure. If the walls presented no elasticity but were rigid, blood pressure would fall to zero between the heartbeats. This fact can very readily be shown by a simple physical model consisting of a pump to represent the heart, connected through a T-piece with two tubes, one of which is elastic, the other rigid. The free end of each tube is contracted to a narrow aperture representing the peripheral resistance, and either tube may be shut off from the pump by means of a stopcock (see Fig. 30). Each tube should also be connected with a mercury manometer. If now the stopcocks are arranged so that the fluid passes into the rigid tube while the pump is in action, it will be found that with each stroke of the pump the pressure in the tube rises considerably, but that it falls to zero between the strokes. If now the stopcocks are turned so that the flow is through the elastic tube, the action of the pump being

meanwhile kept up, it will be found that the pressure between the strokes is maintained at a height which is dependent on: (1) the rate at which the pump is operating, and (2) the resistance to outflow from the tube. The quicker the action of the pump and the higher the resistance, the lower the fall of pressure between the beats.

The physical explanation of this result is clearly that the fluid within the elastic tube when the wave of pressure travels into it from the pump distends the walls of the tube, so that when the pressure from the pump ceases to act, the stretched elastic walls recoil on the column of fluid and maintain the pressure. We may say that the elastic fibers in the vessel walls store up some of the systolic pressure and then transmit it to the blood during diastole.

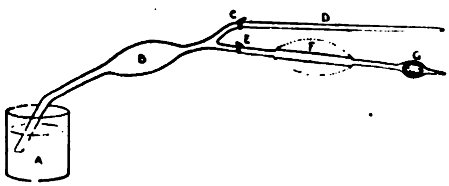


Fig. 31.—Diagram of experiment to show that the diastolic pressure depends on the elasticity of the vessel wall. The pulse (produced by compressing the bulb *B*) disappears when fluid flows through an elastic tube (*F*) when there is resistance (*g*) to the outflow. *A*, basin of water; *B*, bulb syringe; *C* and *E*, stopcocks; *D*, rigid tube; *F*, elastic tube; *G*, bulb filled with sponge.

These considerations would lead us to expect that patients with hardened arteries should exhibit a lower diastolic pressure than normal persons, which, however, is not usually the case, since such patients also suffer from an increase in the resistance to the flow of blood in the periphery. The pressure pulse in these patients is, however, very marked. On the other hand, when the vessel walls become more extensible and elastic, as in certain cases of aneurism, the pressure pulse in the vessels below the aneurism is distinctly less than that observed in normal vessels of the same patient.

CHAPTER XVII

THE ACTION OF THE HEART

Having studied the methods for measurement and the main factors concerned in the maintenance of the arterial blood pressure, we may now proceed to study in greater detail the two most important of these; namely, the action of the heart, and the peripheral resistance.

The heart action has to be studied from two viewpoints, the *physical* and the *physiological.* From the physical viewpoint we have to study the heart as the pump of the circulation. We must see how it acts so as to raise the pressure of the blood within it, and how the valves operate so as to direct the bloodflow always in one direction. We must also explain the causes of certain secondary physical phenomena, such as the heart sounds which accompany the heart action, and of certain secondary changes in pressure produced in the other thoracic viscera by each heartbeat. From the physiologic viewpoint we must investigate the conditions responsible for the constant rhythmic activity of the heart and the control to which this is subjected through the nervous system.

THE PUMPING ACTION OF THE HEART

When the heart is viewed in the opened thorax of an animal kept alive by artificial respiration and lying in the prone position, it can be noted that with each contraction the ventricles become smaller and harder, that the apex tends to rise up a little, so that if the thorax were intact it would press more firmly against the walls, and that it rotates slightly from left to right, but does not move nearer the base of the heart. If the auriculoventricular groove is carefully observed, it will often be noted that it moves slightly toward the apex with each systole, whereas the base of the heart itself, where it is attached to the large vessels, remains fixed. The auricles can often be seen to contract and relax before the ventricles.

The most noteworthy results of this inspection are that during systole the apex of the heart does not move toward the base, but that the auriculoventricular groove moves slightly toward the apex. That these same movements occur in the intact animal can be shown by the very simple experiment of pushing two long steel knitting needles

through the thoracic walls into the heart walls, one of them so placed that it pierces the apex of the ventricle, the other so that it pierces the base. The needles then act as levers with their fulcra at the chest wall, and if the movements of their outer free ends, produced by the movements of the heart, are observed, they will be found to confirm the observations made on the exposed heart.

More particular investigations of the changes occurring in the shape of the heart cavity during systole and diastole have been undertaken by making measurements of sections across the heart in one or other of these conditions. For such purposes the heart in diastole is easily obtained, but for the heart in systole it is necessary to use the somewhat artificial means of injecting the heart with hot chromic acid solution just before the death of the animal. The chromic acid causes the cardiac muscle to contract and maintains it in this condition. The outcome of these investigations is, however, not of much practical importance.

Although it is now common knowledge that the direction of the flow of the blood is from the veins to the arteries, yet it may be of interest to consider for a moment the general principle of the methods by which William Harvey succeeded in making this discovery. His evidence was partly anatomic, partly experimental. He pointed out that the walls of the veins, and of the auricles to which they lead, are very thin, whereas those of the arteries and ventricles are very thick, and he concluded that in the veins the blood must flow gently from the tissues toward the heart, to which the valves in the veins direct it, and that in the arteries it must be propelled by pulses with each systole through the arteries towards the tissues by the contraction of the walls of the ventricles. The experimental support for this hypothesis he furnished partly by clamping the large vessels, veins and arteries leading to or from the heart, and observing the resulting distension or collapse of the vessel; and partly by calculation of the amount of blood which must be expelled from the ventricles in a given period of time.

Harvey's discoveries concerning the events of the cardiac cycle were not much added to until experimental methods were devised by which the pressure changes occurring in the various cavities could be measured and compared. Until such measurements were elaborated, it was impossible to investigate the mechanism by which the various valves between the heart cavities and the vessels connected with them perform their function, or to describe with any degree of accuracy the events occurring in the heart chambers during the various phases of the cardiac cycle. It is for the purpose of ascertaining the exact time relationship of these changes that intracardiac pressure curves are studied.

Intracardiac Pressure Curves

The earliest method for taking such curves consisted in introducing into the cardiac chambers and the blood vessels of the horse, so-called *cardiac sounds.* These consisted of a more or less rigid tube furnished at one end with a little elastic bag or ampulla and connected at the other with a tambour, by means of rubber tubing. One of these little bags was placed in one of the ventricles, another in the auricle or aorta, the tube being inserted in the former case through one of the large veins at the root of the neck; in the latter case through the carotid artery. The intracardiac pressure curves obtained in this way marked a great advance over the methods that had previously been used to study the events of the cardiac cycle, but they were so faulty in comparison with tracings

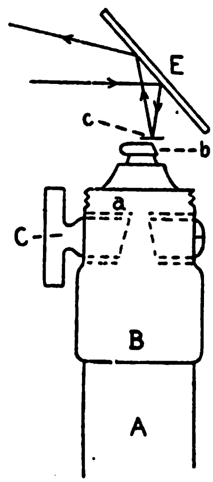


Fig. 32.—Diagram of Wiggers' optical manometer. The wide glass tube (*A*) (connected with the ventricle, etc.) is connected with a brass cylinder (*B*) provided with a stopcock (*C*), the lumen of which comes in apposition with a plate (*a*) having a small opening in it. The freedom of communication between *B* and *a* is regulated by the position of the tap. Above *a* is a segment capsule (*b*) 3 mm. in diameter and covered by rubber dam. This carries a small mirror (*C*) fastened so that it pivots on the chord side of the capsule. Above the capsule is arranged an inclined mirror, from which a strong beam of light is reflected on to the mirror (*c*) on the capsule. This beam then travels back and the mirror (*E*) is adjusted so that it impinges on a moving photographic plate. The slightest movements of the small mirror (*C*) are thus greatly magnified.

taken by more modern methods that it is not worth while considering them any further here.

The physical errors involved in the use of the older instruments were due mainly to the elastic recoil of the membranes, etc., used in their construction. A great improvement in technic was afforded by the use of the spring manometer of Hürthle (see page 126), which was connected with one of the heart cavities by a cannula filled before insertion with some anticoagulant fluid. The cavity of the tambour was made as small as possible, and either left empty or filled with the anticoagulating fluid.

A searching investigation into the physical principles involved in taking records of sudden changes in pressure by such instruments has, however, shown that considerable errors are incurred, the inertia of the moving mass of fluid in the tubing and the necessity of using levers in order to secure records being responsible for most of them (cf. Wiggers). Their elimination has recently been achieved by using a so-called optical manometer, one of which (Wiggers') is shown in the accompanying figure. It consists of a wide glass tube *A*, connected above with a hollow brass cylinder *B*, provided with a stopcock *C*, the lumen of which tapers from below upward till it assumes the same diameter as an aperture in the segment capsule *b*, above it—that is, a capsule cut away at one end—which is 3 mm. in diameter and covered with rubber dam. By adjustment of this stopcock the pulsations of the fluid in *A* and *B* can be damped to a greater or less extent before they are transmitted into the

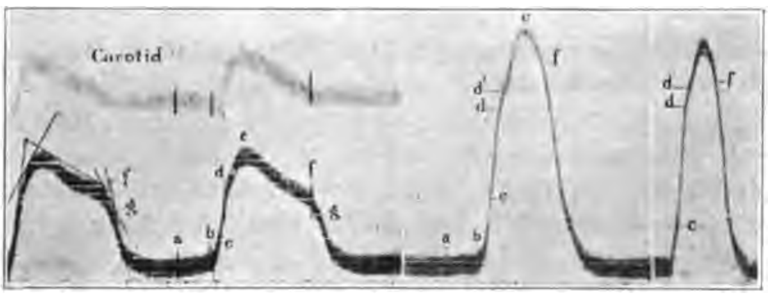


Fig. 33.—Optical records of intraventricular pressure; *a-l*, auricular systole; *b-d*, presphygmic period; *d-f*, sphygmic period; after *f*, diastole. Instruments of varying degrees of sensitiveness were employed in taking the curves. (From Wiggers.)

segment capsule. A small piece of celluloid carrying a tiny mirror rests on the rubber dam, being pivoted on the chord side of the capsule. A mirror is attached to the capsule with its plane so adjusted that the image of a strong light placed at some distance from it is focused on the little mirror carried by the celluloid. The ray reflected from the little mirror and again reflected from the larger mirror is adjusted so as to impinge upon a moving photographic plane travelling at a uniform rate in a suitably constructed photographic apparatus. By the use of such an apparatus the chief errors encountered by the use of the older instruments are eliminated, because there is no moving mass of fluid and there are no levers to set up spurious vibrations. Curves secured by the use of this instrument are shown in Fig. 33.

Two objects must be kept in view in analyzing the curves: (1) Curves obtained from the different cavities may be compared in order to determine the exact moment during the cardiac cycle at which such pres-

sure changes occur as must serve to produce opening or closing of the various valves; and (2) the contour of the curves obtained from each cavity may be examined in order to find out exactly how the pressure in that particular cavity is behaving.

Comparison of the Curves

Before using the curves for ascertaining the relative pressure in the different cavities, they must be graduated according to some scale, for it is clear that by the use of instruments like those we have been describing, the absolute pressure value of each curve will vary according to the construction of the instrument (thickness of membrane, etc.), and indeed instruments of varying degrees of resistance must be employed in taking curves from places having such different pressures as exist in the auricles and ventricles. The graduation is, however, a very easy matter, and consists, as already explained (page 126), in connecting the instrument by means of a T-piece with a mercury manometer and a pressure bottle and then marking on the tracing, the points corresponding to each 10, 20 or 50 millimeters of increase of pressure, as the case may be.

To ascertain the time relationship between the opening and the closing of the auriculoventricular valve, the tracings should be taken from the right auricle and the right ventricle, and to ascertain the same with regard to the semilunar valve, from the left ventricle and the aorta.*

By comparing the curves it is now an easy matter to ascertain the exact moment at which the pressure in the one cavity comes to equal that in the other. This moment, read on the accompanying time tracing, will obviously indicate that at which the particular valve is just about to open or close. From the results of such experiments, the curves may be superimposed as in Fig. 34.

In the first place let us compare *the curves from the right auricle and ventricle.* The curves begin at the very end of diastole, and they show that a distinct increase in pressure is occurring in both auricle and ventricle and lasting about 0.05 second. This is of course caused by auricular systole, and since it occurs in both cavities, it indicates that the passage between them, the auriculoventricular orifice, must be open. The ventricular curve then suddenly shoots away beyond the auricular because of the onset of systole in the ventricle, and the point at which the two curves begin to separate indicates the moment at which the auriculoventricular valves close. From this time on until ventricular systole has given place to diastole, (about 0.2 second), the auricle is

*The connections with the heart may be made by pushing long cannulæ down the large veins or arteries, or in the case of the ventricles by inserting a cannula with a sharp point directly through the wall of the ventricle.

therefore shut off from the ventricle. The exact moment in diastole at which the two cavities are again brought into communication—i.e., the ventricular valves open—is indicated by the curves coming together.

Having thus determined the exact moments of opening and closing of the auriculoventricular valve, we may now proceed to compare the intraventricular pressure curve with that taken from the aorta. After the necessary calibration corrections, this curve has been placed in Fig. 34 in its true relationship to the ventricular curve. Beginning again at the end of diastole, we find that the aortic pressure is very considerably above that of the ventricles, indicating that the semilunar valves must be closed; and it will be observed that the intraventricular pressure at

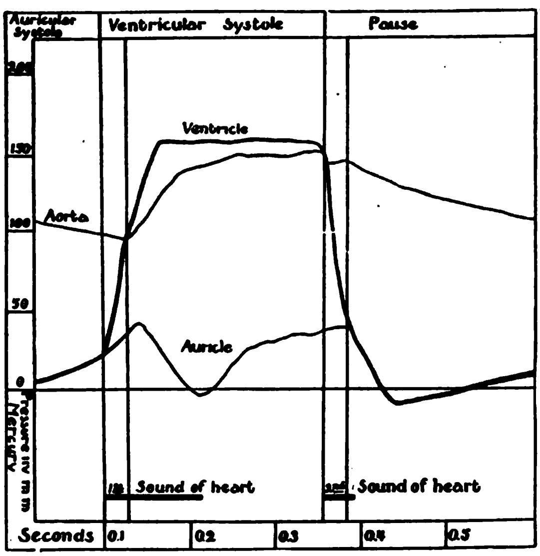


Fig. 34.—Pressure curves after being graduated have been superimposed. The presphygmic, sphygmic and postsphygmic periods of ventricular systole are shown by the vertical lines. The *A-V* valves close at the first line. The aortic valves open at the second line and close again at the third line. The *A-V* valves open at the fourth line. The position of the two main heart sounds is also indicated.

the beginning of systole does not rise sufficiently to open them until an appreciable interval (0.02 to 0.04 second) after the closure of the auriculoventricular valves; that is to say, there is a period at the beginning of ventricular systole during which the ventricle is a closed cavity. It is a period during which the ventricle by its contraction is getting up a sufficient amount of pressure in the fluid contained in it to force open the semilunar valves against the resistance of the pressure in the aorta, and it has been popularly called "the period of getting up steam," or, in physiologic language, the *isometric,* or the *presphygmic,* period. We shall use the last-mentioned term in our further discussion here.

After the aortic valves have been opened, it will be observed that the pressure in the ventricles is just a little above that in the aorta, and that it continues so during the whole of ventricular systole. When diastole sets in, the pressure in the ventricles quickly falls, and a point is soon reached at which equality of pressure in ventricle and aorta is again attained. This corresponds to the moment of the closure of the semilunar valves. The pressure in the ventricle, although now rapidly falling, takes a little time before it has fallen low enough to permit the auricular valves to open. Here again, then, the ventricle is a closed cavity, and we have what is known as *the postsphygmic period.*

CHAPTER XVIII

THE PUMPING ACTION OF THE HEART (Cont'd)

THE CONTOUR OF THE INTRACARDIAC CURVES

The Ventricular Curve

From an analysis of the contour of each curve, further interesting points are brought to light. The ventricular curve in the diagram alluded to above (Fig. 34) is shown as having a flat top or plateau. By the use of the more modern, optically recording, instruments it has been shown that this plateau becomes displaced by a peak if every precaution is taken to prevent dulling down of the pressure changes in the instrument, as by opening wide the stopcock in the instrument (Fig. 33). The peak is, however, by no means a sharp one, so that we may fitly describe the contour of the ventricular curve during the sphygmic period as consisting of a *rising portion,* almost continuous with the curve during the presphygmic period, a *summit* and then a *declining portion,* which is usually slower than the ascending. The practical value arising from a study of the curves lies in the insight which they give us into the nature of the stroke of the cardiac pump. They show us that the impulse which the ventricle gives to the moving mass of blood in the aorta is rather a sudden than a sustained one. The column of blood in the aorta is a mighty thing to move, and it would appear as if a sustained pressure brought to bear on it during the sphygmic period would be far more efficient in bringing about an adequate movement of the blood than a sudden jerk. In closing a heavy gate a slow sustained pressure is far more effective than a short push.

It is further of interest to note on the intraventricular pressure curve that there is very little indication of any *secondary waves or vibrations* at the moment during which the semilunar valves are opened or closed. Nevertheless, by close scrutiny it can usually be seen that a slight change in the direction of the ascending curve is evident when the valves open (see Fig. 33), and similarly that the moment of closing is indicated by a sharper bend in the curve. As a matter of fact, Wiggers has shown that the exact contour of the curve during the sphygmic period depends partly on the degree of sensitiveness of the optical manometer used and partly on the tension existing in the ventricle just before contraction.

In the case of the right ventricle the contour of the curve also depends on the degree of resistance to the bloodflow through the pulmonary circuit. The top of the curve becomes broader when the initial tension is high, and more rounded when there is a high pulmonary resistance.

Another point of interest in connection with the ventricular curve is that early *in diastole it descends below the line of zero pressure,* indicating that a negative or suction pressure must exist in the ventricle at this time. It will be further observed, however, that this subatmospheric pressure exists for only a very short time. The auriculoventricular valves being opened, a similar negative pressure is also present in the auricular tracing. Were we to depend on such records alone for evidence of the actual existence of this negative pressure in the heart, objection might be taken to the conclusion on the ground that it was due to the

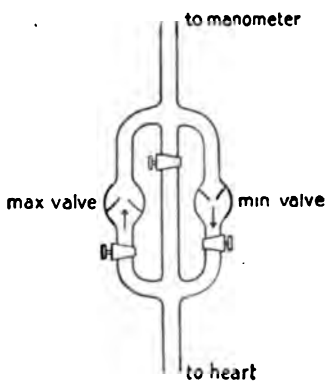


Fig. 35.—Von Frank's maximal and minimal valve, which is placed in the course of the tube between heart and mercury manometer. By turning the stopcocks, it may be used as a maximum, minimum, or ordinary manometer (central tubes open). (From Starling.)

sudden recoil to which the instrument is subjected at the beginning of diastole. It is necessary therefore to control these observations by the use of an entirely different method. This consists in connecting the heart with a *valved mercury manometer* (see Fig. 35). This instrument does not of course record any sudden changes of pressure in the cardiac cavity, but in obedience to changes in pressure the mercury slowly moves in the direction in which the valve permits it to move. Such an instrument, with the valve opening towards the heart, is called a minimal manometer, and after it has been connected with the ventricle, it will be found that a negative pressure of perhaps 40 or 60 mm. Hg is recorded.

Evidently, then, *the negative pressure does actually exist in the ventricle* during some phase of the cycle, and the question arises as to whether it is of importance in connection with the pumping action of the heart. At first sight, considering the heart as an elastic structure, we might con-

ceive that the negative pressure would serve to suck blood into the heart, just as it sucks water in an ordinary ball syringe. Closer consideration will, however, show that this conclusion is untenable, partly because the negative pressure exists in the ventricle for so short a period of time, and partly because it would have to operate on the slowly moving column of blood in the thin-walled veins, with the result that it would cause the walls of these vessels to come together rather than produce a movement of the blood contained in them. The negative pressure of the heart can not therefore be of much consequence in attracting the venous blood into the ventricle.

Several factors may cooperate to produce this negative pressure, among them being the sudden opening out of the base of the ventricles at the beginning of diastole, the elastic recoil of the tissue which becomes compressed in the heart walls during systole, the turgescence of the walls of the ventricles produced by the sudden inrush of blood into the coronary vessels at the beginning of diastole, all of which factors tend to cause an opening out of the walls of the ventricles with a consequent increase in the capacity of their cavities.

The Auricular Curve

Examination of the intraauricular pressure curve is of particular interest because of the relationship which it has to a tracing taken of the movements in the jugular vein at the root of the neck (see page 274). This jugular pulse curve, as it is called, is produced mainly by the changes of pressure occurring in the auricle, from which it differs only in the relative height of the various waves. By graduating the intraauricular pressure curve by the method described above, we can tell exactly the magnitude in the changes of pressure occurring during each cardiac cycle. This obviously can not be done with a tracing taken from the jugular vein, although qualitatively the tracings reflect exactly the changes that are occurring in the auricle.

On examining the auricular pressure curve (consult Figs. 34 and 97), we find that after the wave of presystole, which of course coincides exactly with that on the intraventricular curve, a second wave occurs culminating in a peak almost exactly at the beginning of the sphygmic period. The curve then rapidly descends, usually indeed below the line of zero pressure, and slowly rises throughout the rest of ventricular systole, until the moment of opening of the auriculoventricular valve, when it descends again and thereafter runs parallel with the ventricular curve. The letters used to designate the waves are the same as those employed for similar waves shown on the jugular pulse tracing, and although the

lettering is more or less arbitrary, we must accept it because of its general usage in all work of this kind.

As to the causes of the waves *A* is of course caused by auricular systole or presystole; *C*, occurring as it does at the beginning of the period of ventricular systole, is caused by the bulging into the auricle of the closed auriculoventricular valve. The floor of the auricle, in other words, at this moment becomes somewhat elevated and imparts to the blood which is resting upon it a slight wave of pressure, which is transmitted along the veins for a considerable distance. The succeeding depression is marked *x*, and the negative pressure which it indicates is probably due to the co-operation of three forces, all tending to increase the auricular capacity: (1) the diastole of the walls of the auricle; (2) the descent of the auriculoventricular groove, thus tending to open out somewhat the folds in the walls of the auricle; and (3), probably most important of all, the tendency of the thin-walled auricles to become dilated as a result of the sudden diminution in intrathoracic pressure produced at each heart-beat by the discharge of blood from the heart and intrathoracic blood vessels into those of the rest of the body. All thin-walled structures in the thoracic cavity, the auricles included, will expand to take up the extra room created in the thoracic cavity. Similar negative heart pulses, as they are called, can be observed with each systole in the lungs and in the esophagus.

THE MECHANISM OF OPENING AND CLOSING OF THE VALVES

When physical valves open and close as a result of the changes in pressure on their two surfaces, a certain amount of fluid must succeed in passing the valve flaps before these become perfectly closed. But there is every reason to believe that such is not the case in the heart, the flaps of both the auriculoventricular and the semilunar valves being already completely closed before pressure conditions entailing a possible regurgitation of blood through them become established.

Auriculoventricular Valves

During diastole the flaps of the auriculoventricular valves are hanging down into the ventricle and floating in a half-open position in the blood, which is meanwhile accumulating in the chamber. This position is dependent upon the operation of two opposing forces on the valve flaps: the pressure of the blood flowing from the auricle on their upper aspects, and reflected waves of pressure from the walls of the ventricle on their under aspects (centripetal reflux). When presystole occurs, the pressure of the auricular stream momentarily increases, thus slightly distending the wall of the meanwhile relaxed ventricle and after a moment's

delay causing the reflected wave to become more pronounced. At the same time the muscular fibers in the valve flaps (Kürschner's fibers) contract and make the flaps shorter, the total effect of the two factors being that the valve takes up a position nearer that of closure. When presystole suddenly stops, the reflexed waves will persist for an instant of time longer than the auricular wave which causes them, because of the elastic nature of the ventricular wall, so that the valve flaps close with perfect opposition not merely at their edges but also for a considerable distance along their upper surfaces.

When ventricular systole starts, the only effect of the high pressure which is brought suddenly to bear on the under surfaces of the already closed valves is to cause them to vibrate and to bulge into the auricles, being meanwhile anchored down and prevented from flapping into the auricle by the chordæ tendineæ. Although there is reason to believe that the musculi papillares to which these are attached begin to contract at the

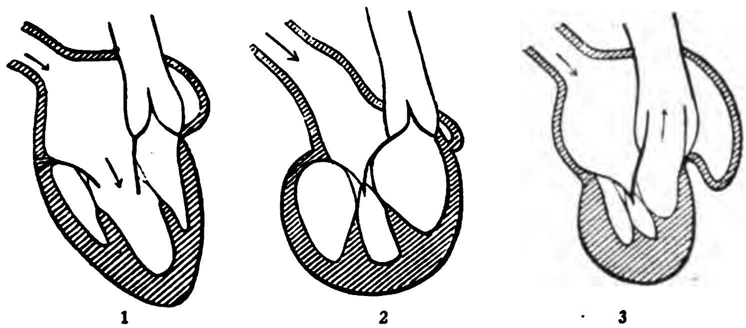


Fig. 36.—Diagram to show the positions of the cardiac valves: 1, during diastole; 2, during the presphygmic period; 3, during the sphygmic period.

very outset of ventricular systole—indeed slightly to precede it (see page 263), and thus keep the chordæ taut, yet as systole continues the contraction of these muscles becomes more and more pronounced, and the resulting tightening of the chordæ serves to draw down the valve flaps, so that progressively larger proportions of their upper aspects tend to become opposed. Meanwhile the auriculoventricular orifice is also becoming narrowed down on account of the contraction of the musculature of the auriculoventricular groove.

Semilunar Valves

The mechanism involved in the operation of the semilunar valves is somewhat different. It has been shown that, when fluid is flowing in a tube, the pressure and velocity are not equal in the axial and peripheral parts of the stream. In the axis the velocity is greater than in the layers of fluid next to the walls, but the pressure is less. These facts can be

demonstrated by observing the flow through a wide tube of water in which are suspended lycopodium spores. By placing in the tube small bent tubes so arranged that one open end lies near the periphery and the other near the center, it can be seen that the differences in pressure are such as to cause the fluid to flow from periphery to axis (centripetal eddies).

If the bent tubes are used to study the conditions of flow in a tube which suddenly becomes wider, it will be found that where the wide portion starts centripetal eddies are set up, which tend to carry the seeds into the axis of the stream, where their velocity is greatly increased. Now these are the conditions obtaining at the beginning of the large arteries

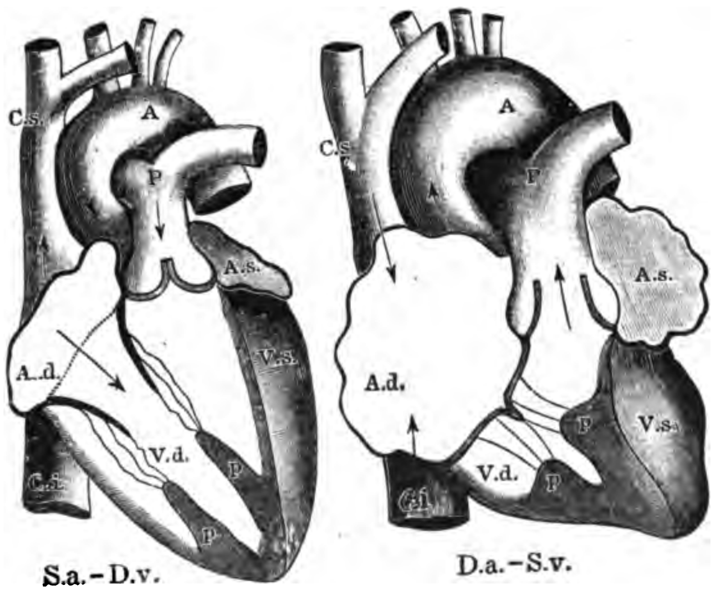


Fig. 37.—Diagram showing the position of the cardiac chambers and valves during presystole (S.a.—D.v.) and during the sphygmic period. (From Landois.)

of the heart, the orifice into the ventricles being constricted, while at the sinus valsalvæ the vessels are dilated. A centripetal vortex must be set up in the sinus, tending to throw the valve flaps into a closed position, which, however is prevented by the blood rushing between them from the ventricles. They thus take up a mid-position and vibrate in the stream. When the efflux from the ventricle stops at the end of systole, the reflux, lasting for a moment longer and being now unopposed, immediately closes the valves, in which position they are then maintained by the greater pressure on their upper surfaces.

The position of the valves relative to the events of the cardiac cycle is shown in Figs. 36 and 37.

THE HEART SOUNDS

During certain phases of the cycle distinct sounds, the heart sounds, can be heard by applying a stethoscope to the thoracic wall. The first occurs at the beginning of ventricular systole and is best heard over the apex beat; the second occurs at the beginning of diastole and is heard best at the second right costal cartilage or in the second left intercostal space. A third sound, much less distinct, is sometimes heard in diastole a short time after the second. To study the exact time relationship of the sounds the vibrations which they set up can be recorded graphically alongside cardiac tracings by means of a microphone attachment to the electrocardiograph (see page 259).

Causes of Sounds

It has been found that the *first sound* consists of two distinct elements, one high pitched and the other of a dull character. The former element is believed to be the result of vibrations set up in the flaps of the auriculoventricular valves, and therefore in the blood in the heart, by the sudden rise in systolic pressure. The dull element on the other hand is undoubtedly of muscular origin. The evidence for these conclusions is as follows: (1) When the auriculoventricular valves are prevented from closing properly either by disease or by pushing a loop of wire down the large veins, the high pitched quality disappears, and nothing but a rushing sound accompanies the dull bruit produced by the contracting muscle. (2) In a heart that has been rendered bloodless by an incision near the apex, or even in an excised but still beating heart, the dull element of the first sound still continues to be heard for a short time. That contracting muscle produces a sound is a well-established fact.

There are, however, many obscure phenomena connected with the causation of the first sound, but we can not go into such controversial matters here. A close inspection of the electrophonographic tracing shows that the sound starts at the beginning of the presphygmic period, and that it lasts with gradually declining variation in intensity until well into the sphygmic period (Fig. 38).

The *second sound* occurs accurately at the beginning of diastole and can readily be shown to be caused by the sudden shutting and stretching of the semilunar valves, which throws them, the blood in contact with them, and the neighboring walls of the aorta into vibration. Proof of this conclusion is furnished by the following facts: The second sound immediately disappears if the blood is let out of the heart by opening the apex, and it is replaced by a rushing "bruit" if the flaps are prevented from closing as a result of disease or of hooking them back by

passing a wire down the carotid artery. The *third sound,* although audible only in some individuals, can nevertheless be shown to exist by the electrophonograph, and since it occurs at the time when the auriculoventricular valves open, it is believed to depend upon the sudden inrush of blood from auricles to ventricles.

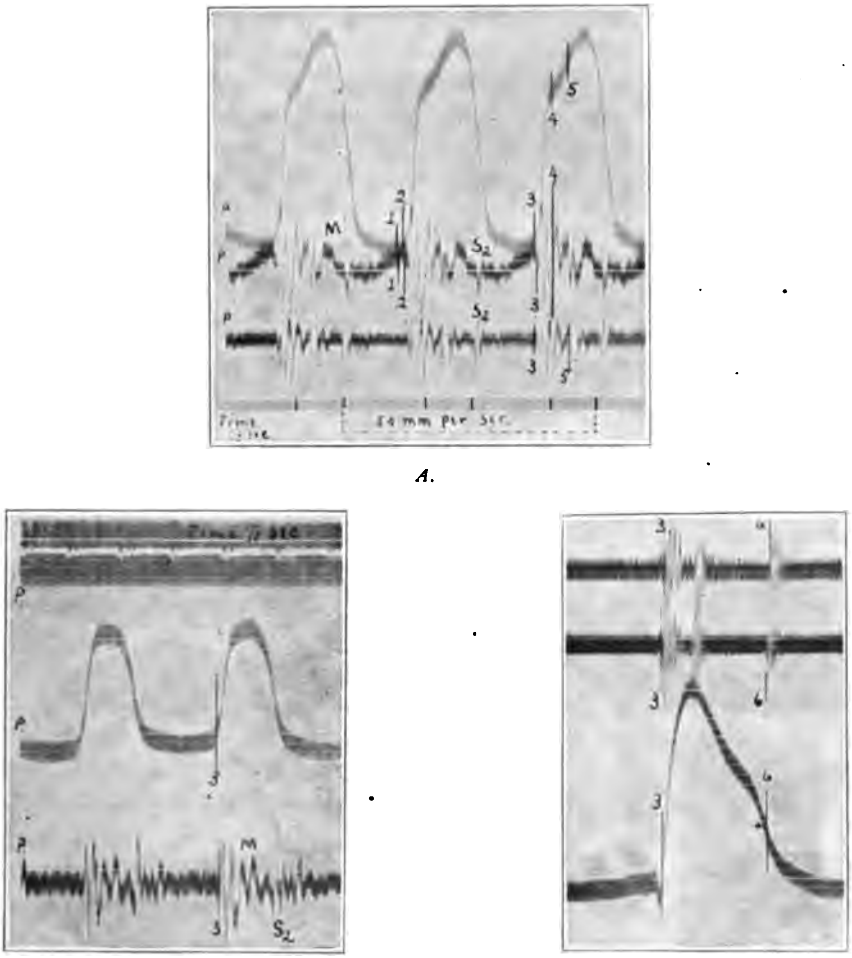
The greatest importance of the sounds is in the clinical diagnosis of valvular and other lesions of the heart. When a valve leaks, for example, the blood escapes past it under great pressure, and is ejected into a mass of blood at low pressure, these being conditions which are well known to create sounds or *bruits.* By examining the exact relationship of such bruits to the normal heart sounds, deductions can be drawn concerning the condition of the various valves.

Record of Heart Sounds

The heart sounds have been graphically recorded by transmitting them through a stethoscope to a microphone placed in circuit with a string galvanometer (electrophonograms). Through this circuit passes a current the strength of which depends on the resistance offered by the microphone, this resistance being proportional to the number and amplitude of the vibrations of the sounds transmitted to it through the stethoscope. There are several objections to this method. One of these is dependent on the varying distance of the heart from the chest wall, which causes many of the sound vibrations to be lost before they reach the stethoscope; another, on adventitious sounds arising from contracting muscles, the impact of the heart against the chest wall, etc., and still another on unequal resonation by the air in the neighboring portions of lungs. To investigate the problem more thoroughly, Wiggers,[37] using anesthetized animals, has recorded the sounds by carefully stitching to the heart (exposed through a small opening in the pericardium) a lever, the end of which was attached to a "transmitter" consisting of a small capsule covered with rubber dam. The transmitter was connected by rubber tubing to a "recorder" consisting of another small capsule carrying on its membrane (made of rubber cement) an eccentrically placed small mirror, on to which a beam of light was thrown. The movements of the beam of light reflected from the mirror, and caused by the sound vibrations, were photographed. Mechanical vibrations set up in the apparatus itself were largely eliminated by a side opening on the recorder, and the effect of outside sounds minimized by surrounding the recorder by a ventilated glass housing.

Although this apparatus is not free from faults due to inherent vibration frequency and resonance, the records secured by it are valuable in showing the exact relationship of the sounds to the events of the cardiac

cycle. The vibrations from the two ventricles are alike, but differ from those taken from the aorta. *The first ventricular sound* consists of from five to thirteen irregular vibrations, usually in three groups, the first composed of two small vibrations, the middle one of several large vibrations, and the third of a varying number of small vibrations. The



A. *B.* *C.*

Fig. 38.—Electrophonograms along with intraventricular pressure curves from three different experiments. In *A* the uppermost curve shows the pressure, the middle one the sounds of the right ventricle, and the lowermost one those of the aorta. *P* indicates the relative position of the curves. *M* is due to mechanical oscillations. *S2* indicates the second sound, and *1, 2, 3,* and *4* the corrected time relations of the first sounds. In *B,* the pressure and sound curves are both from the left ventricle (letters same as in *A*). In *C,* the aortic and pulmonary arterial sounds are shown (letters same as in *A*). (From Wiggers and Dean.)

duration of the sound is from 0.05 to 0.152 seconds, and the periodicity from 0.004 to 0.054 per second. When compared with an intraventricu-

lar pressure curve, the initial vibrations occur 0.01 second prior to the rise in pressure, the main vibrations reaching their greatest amplitude before the sphygmic period begins, and the final vibrations occurring during the early part of the sphygmic period and therefore just before the aortic pressure has reached its height. The main vibrations therefore occur during the descending limb of the *R* wave of the electrocardiogram (beginning 0.01 second before its completion), the small preliminary vibrations occurring during the ascending limb. When taken from the aorta, the record of the first sound is somewhat different, there being no initial vibrations and the main ones being of greater frequency and reaching their maximum earlier than those taken from the ventricle. The subsequent vibrations are also larger, especially when the aortic pressure is high (Fig. 38).

The record of the *second sound* at the ventricle is much simpler and usually of less amplitude than the first, consisting of two to six vibrations lasting 0.015 to 0.056 second. They begin a short time after the ventricular pressure begins to fall, approximately at the dicrotic notch of the aortic curve, being completed in from 0.015 to 0.025 second after the bottom of the notch. Their relationship to the *T* wave is variable. Taken from the aorta, the record of the second sound shows vibrations of greater amplitude and of a greater frequency than that from the ventricle.

CHAPTER XIX

THE NUTRITION OF THE HEART

THE BLOOD SUPPLY

In cold-blooded animals, such as the frog, the heart muscle is nourished by blood soaking into it from the heart chambers, which indeed do not form definite cavities as in the mammalian heart, but exist as an interlacement of muscular tissue. In the hearts of higher animals, the musculature is supplied by special arteries (the coronary), although a certain amount of blood may still pass directly from the cardiac cavities into the musculature through the veins of Thebesius.

The relative importance of the various branches of the coronary artery in maintaining an adequate nutrition of the heart has been studied by observing the effect of occlusion of one or more of them (W. T. Porter[9].) Occlusion of the circumflex branch of the left coronary artery caused arrest of the heartbeat in about 80 per cent of cases, the arrest being usually accompanied by fibrillary contraction. Occlusion of the right coronary arrested the ventricular contraction in about 20 per cent of the cases. Smaller branches may be occluded without any evident change in the heartbeat.

These results indicate that the capillary areas supplied by the branches of the coronary artery do not freely anastomose with one another. They are more or less terminal arteries; that is, each branch supplies a distinct region of the cardiac muscle. If one of the smaller branches of the coronary is occluded, although there is no immediate stoppage of the heartbeat, yet after some time the area supplied by that branch usually undergoes necrosis, again indicating that collateral circulation can not have become established. It is interesting, however, to note in this connection that anatomic studies have shown that a certain amount of anastomosis does occur between capillaries of different branches, although it is evident, from the above observations, that no adequate collateral circulation becomes established through this anastomosis.

PERFUSION OF HEART OUTSIDE THE BODY

In order that the blood supply through the coronary arteries may adequately maintain the normal nutrition of the cardiac muscle, certain

conditions must be fulfilled. The recognition of these conditions has been accomplished by observations on the excised heart, for it has been found that if they are fulfilled the mammalian heart can be made to beat in perfectly normal fashion for several hours after its removal from the animal's body. Indeed certain mammalian hearts, such as that of the rabbit, may be made to beat for several days outside the body. We may consider the essential conditions of the blood supply under four headings: (1) the temperature; (2) the oxygen supply; (3) the pressure; and (4) the chemical composition. Successful perfusion may be performed with artificial saline solutions (e. g., Locke's), but it is simplest in investigating the relative importance of the above conditions to start the heart perfusion with defibrinated blood.

After bleeding an anesthetized animal, such as a dog or a cat, until no more blood can be removed, the blood is defibrinated and filtered through gauze to remove the fibrin. The thorax of the dead animal is then quickly opened, ligatures placed around the main arteries springing from the arch of the aorta, a cannula with its end pointing toward the heart inserted into the descending thoracic aorta, and the latter cut across below the point of insertion of the cannula. The heart is then quickly removed from the thorax and an artificial saline solution (Locke's) allowed to run into the aortic cannula through a side tube, until all the blood has been washed out from the coronary vessels. During this operation the heart may develop a few beats even though the solution is quite cool. The aortic cannula is now connected with a bottle containing the defibrinated blood diluted with Locke's solution and brought to body temperature by immersion in a water-bath. By means of a suitably regulated air pressure exerted on the surface of the diluted blood in the bottle, this is forced through an outlet at the foot of the bottle into tubing which runs to the aortic cannula. The fluid thus finds its way into the coronary vessels; for in passing toward the heart in the aorta it will close the semilunar valves and force its way under pressure into the coronary vessels, subsequently escaping by the coronary sinus into the right auricle. Very soon after the perfusion is started the heart begins to beat vigorously and regularly, thus offering a suitable preparation upon which to test the first three mentioned conditions necessary for the nutrition of the cardiac musculature (Fig. 39).

If the *temperature* of the solution is allowed to fall considerably, the beat becomes much slower, and if the cooling is proceeded with, the heart will after a while cease beating altogether. If the *pressure* is lowered, the beat will not necessarily become slower but very much feebler, and will soon cease. In general it may be said that the temperature of the solution affects the rate of the beat, and the pressure affects its strength.

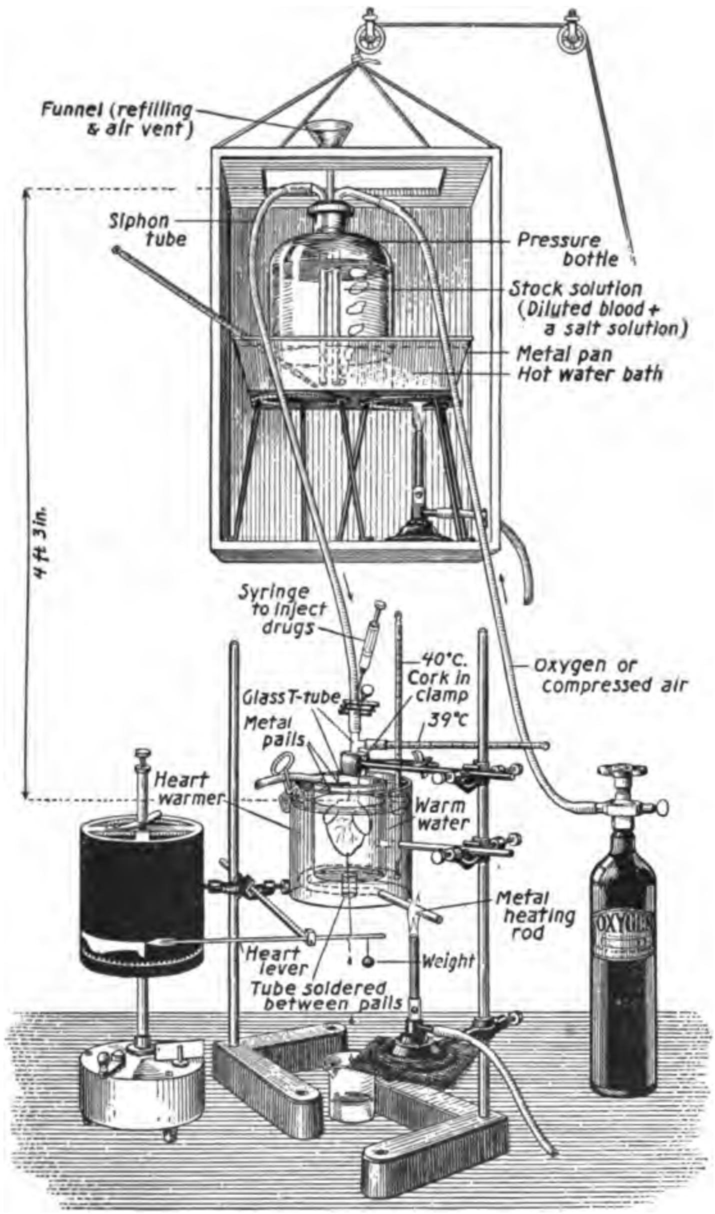


Fig. 39.—One form of apparatus for recording tracings from an excised heart (Langendorff method). The heart is kept warm by a water bath (heart warmer), and the perfusion fluid is also warmed. The driving pressure in this apparatus is supplied by gravity. (From Jackson.)

It is, however, obvious that in perfused preparations changes in pressure are likely to cause alterations in rate as well as in force, unless great care is taken to keep the heart itself as warm as the perfusion fluid.

The importance of an adequate pressure in the coronary vessels has been clearly brought out in certain experiments in which the beat has been maintained for a short time by establishing a pressure in the coronary vessels by means of indifferent fluids or gases. Thus, if oxygen gas is allowed to pass through the vessels under pressure, the heart will beat for a short time, and the same result has been observed even when mineral oil or mercury has been perfused under pressure (Sollmann).

The necessity for an *adequate oxygen supply* is very readily demonstrated. If the darker blood ejected from the right auricle with each heartbeat is transferred immediately to the perfusion bottle, the heartbeat will soon become feeble and irregular, to be readily restored to normal when this dark blood is shaken up with air or oxygen.

By artificial perfusion in the manner above described, the automatism of the heart may be restored many hours after death. Partial restoration, confined to the auricles or to that part of the ventricles lying immediately adjacent to the large blood vessels, can also be accomplished in the heart of man several days after death, provided death has not been caused by some acute toxic infection such as diphtheria or septicemia. The Russian physiologist Kuliabko, has succeeded in restoring for over an hour the normal beat of the heart of a three-months-old boy twenty hours after death from double pneumonia, but here again the pulsation returns only in certain parts of the heart. As will be pointed out, the remarkable resistance of the heart muscle displayed in these experiments has been taken as an argument in favor of the myogenic hypothesis for automatic rhythmic power of cardiac muscle, the argument being that nervous structures could not live so long a time after death. The fallacies in this argument are discussed elsewhere.

RESUSCITATION OF THE HEART IN SITU

A suitable intracoronary pressure is a *sine qua non* for the maintenance of the heartbeat, and this is a fact of great clinical significance, for it indicates that any attempts to resuscitate a dead animal are certain of failure unless the method is such as will bring a nutrient fluid under a certain pressure to bear on the coronary arteries. Injection of fluid, even of defibrinated blood, into a *vein* will obviously fail to fulfill this condition, for the perfusion must be made into an artery so that the fluid is carried down the aorta and thence into the coronary arteries.

The practical question, in so far as resuscitation of the heartbeat is concerned, is therefore, *How can we get the necessary fluid under pressure into the beginning of the aorta?* Even if we were to transfuse fluid under considerable pressure into the aorta through the carotid artery, it would mainly follow the large vessels leading away from the heart. only a fraction of it reaching the beginning of the aorta. To compel the fluid to pass towards the heart we must introduce some obstruction to its passage peripherally. This can be done by the injection of a considerable dose of epinephrine (adrenaline) in normal saline solution through the needle of a hypodermic syringe inserted into the tubing leading from the burette or pressure bottle to the cannula in the carotid artery. As the perfusion fluid is running in, the epinephrine injection is quickly made, artificial respiration and cardiac massage being meanwhile practiced. In the majority of animals it will be found that complete restoration of the normal blood pressure can be effected by this method. Indeed by performing the resuscitation under aseptic conditions, some animals may be permanently resuscitated so far as the circulation is concerned, although the nervous structures, even after a few minutes of "death," never reacquire their normal condition.

The epinephrine acts mainly by constricting the small arterioles and thus directing the bloodflow towards the heart, but partly also by a direct stimulating action on the cardiac muscle. It does not, however, contract the coronary vessels; on the contrary, it is said to cause these slightly to dilate.

THE RELATIVE IMPORTANCE OF THE VARIOUS CONSTITUENTS OF THE PERFUSION FLUID

We can study the chemical conditions necessary for resuscitation of the heartbeat by observing the beat of an artificially perfused heart while solutions of different chemical composition are being perfused through the coronary vessels. At the outset we are impressed with the fact that for successful resuscitation the organic constituents of the nutrient fluid are of trivial importance compared with the inorganic constituents. With a solution containing the proper proportion of inorganic salts, and of course an adequate supply of oxygen, the heart of a rabbit, for example, may be made to continue beating for several days. It is true that it will beat longer if some of the organic constituents of the blood plasma, particularly carbohydrate, are present, but on the inorganic constituents alone its ability to beat is truly remarkable.

Observations on Cold-Blooded Heart

The earlier experiments for the investigation of the chemical conditions necessary for the maintenance of the heartbeat were performed on *the heart of the frog or turtle.* By perfusing either of these hearts with physiologic sodium-chloride solution, it was observed that though the beat might continue for some time, yet it gradually grew feebler and feebler, until at last it ceased altogether with the heart muscle in a condition of extreme relaxation or diastole. If small proportions of potassium and calcium salts (as chloride) were added to the sodium-chloride solution, the beat was much better maintained. Doctor Sidney Ringer proved that the optimum concentration to produce efficient and prolonged contraction for the heart of the frog or terrapin is as follows: potassium chloride, 0.03 per cent; calcium chloride, 0.025 per cent. The effectiveness of the solution was also found to be increased by the addition of 0.003 per cent of sodium bicarbonate. This acts as a buffer substance (page 36), holding the hydrogen-ion concentration at a constant level. More recent work has shown that the hydrogen-ion concentration of the perfusion solutions is of considerable importance in determining the efficiency of the beat, but the optimum is not the same for the hearts of different kinds of animal, and indeed it may differ for different parts of the same heart.

The question naturally arises as to the relative importance of each of the above salts; or rather, we should say, cations, since the anion, chlorine, is the same for all of them. The function of the *sodium* chloride in the solutions is twofold: (1) to endow the solution with the proper osmotic pressure (see page 4); and (2) to perform the special role of the sodium ion in the origination and maintenance of the automatic beat. The latter function of Na can be shown by observing the behavior of strips cut out from the ventricle of the turtle heart and placed in solutions of correct osmotic pressure but containing no sodium chloride—isotonic solutions of cane sugar, for example. They soon cease to beat, but if a small amount of sodium chloride is added to the cane sugar solution, rhythmic contractions return. The role of the *calcium ions* is almost entirely a pharmacologic one. If a strip of turtle ventricle which has been made to cease beating by immersion in isotonic sugar solution is placed in a weak solution of calcium chloride before it is transferred to sodium chloride solution, the spontaneous contractions will return earlier and continue for a longer time. On the other hand, if more than the correct amount of calcium salt is present in the solution, the beats will soon be found to become smaller and smaller in amplitude, because relaxation does not properly occur between them, and ultimately they will cease altogether with the ventricle in a condition

of extreme contraction, called calcium rigor. The importance of *calcium* may also be shown by attempting to perfuse a turtle heart with blood serum from which the calcium has been removed by the addition of sodium oxalate (which precipitates it as insoluble calcium oxalate). The heart soon ceases to beat, but can readily be made to do so again by adding a slight excess of calcium chloride.

The *potassium ions* do not appear, like those of calcium and sodium, to be absolutely essential for the maintenance of the heartbeat; at least the heart of the turtle will beat for a long time when perfused with a solution containing only sodium and calcium salts. The explanation of this result need not, however, necessarily be that potassium is an unessential constituent of the perfusion fluid, for it may well depend on the fact that there is a sufficient store of potassium locked away in the muscle fiber to supply the requirements of the heart muscle for this ion for at least as long as the beat would continue under any circumstances. In any case, we know that potassium has a profound influence on the heartbeat, for when the proportion of it in the perfusion fluid is increased, the beat becomes very slow and the tone of the heart is greatly diminished—that is, it becomes extremely relaxed between the beats; and if the amount is further increased, will very soon come to a standstill in a greatly dilated or diastolic position.

The striking antagonism displayed by these inorganic cations upon the heartbeat has led some investigators to suggest that the stimulus responsible for the rhythmic activity of the heart depends on some sort of chemical union occurring between the inorganic cations and the contractile substance of the heart. Union of calcium with the contractile substance will lead to systole or contraction, whereas union of sodium or potassium will lead to relaxation or diastole.

Observations on Mammalian Heart

Investigation of the efficiency of various saline solutions on the isolated *mammalian heart* has shown that the proportion of the above salts must be somewhat different from that used for the cold-blooded heart. As might be expected, the most efficient proportions are those present in the blood serum of the particular animal whose heart is being perfused. Basing his proportions upon the results of analyses of the inorganic constituents of mammalian blood serum, Locke found that an inorganic solution of the following composition is most efficient: sodium chloride, 0.9 per cent; calcium chloride, 0.024 per cent; potassium chloride, 0.042 per cent; and sodium bicarbonate, 0.01 to 0.03 per cent. When "Locke's solution," as it is called, is perfused, with oxygen in it, under pressure through the isolated mammalian heart at body tempera-

ture, efficient beating can be maintained for many hours. More recently a solution known as Tyrode's is commonly used. It contains a small amount of magnesium and of phosphates. Although undoubtedly superior for some perfused preparations, such as the intestine, it does not seem to be in any way superior to Locke's for the perfusion of the heart. The bicarbonates and phosphates in these solutions endow them with a hydrogen-ion concentration near that of the blood (slightly on the alkaline side of neutrality), and at the same time they act as buffer substances.

As already pointed out, the organic constituents of such perfusion fluids do not appear to be relatively of nearly so much importance as the inorganic. Nevertheless it appears that a small percentage (0.01 per cent) of glucose does materially improve the nutritive qualities of the solution, and it has moreover been shown that after a while the concentration of glucose in the perfusion fluid distinctly decreases. This does not of itself necessarily mean that the glucose is actually utilized by the heart muscle: it might be stored away in it as glycogen. That some consumption of carbohydrate does however occur in the heart has been demonstrated by measuring the intake of oxygen and the output of carbon dioxide through the lungs of an isolated heart-lung preparation perfused outside the body with defibrinated blood. By experiments of this type the attempt has been made to show that the heart of diabetic animals loses the power of burning glucose as compared with the hearts of normal animals. While the experiments are very suggestive, the results do not as yet justify us in claiming that in the latter disease the power of burning glucose in the tissues has been materially depressed.

The concentration of hydrogen ions in the perfusion fluid has an important influence on cardiac efficiency. We also know that the most convenient method for changing the hydrogen-ion concentration of such fluids is by altering their tension of carbon dioxide (see page 354). In a heart-lung preparation,* such alteration in carbon-dioxide tension can very readily be brought about by altering the percentage of this gas in the air with which the lungs are ventilated. To measure the efficiency of the heartbeat in such an experiment, it is convenient to enclose the organ in a cardioplethysmograph, the tracing of which will tell us the degree to which the heart is contracted or relaxed, as well as the output of blood per minute. By increasing the tension of carbon dioxide, it has been found in such experiments that the dilatation of the ventricle is encouraged, so that the heart with each beat discharges a larger quantity of blood (Fig. 40). When defibrinated blood is used the optimum

*A heart-lung preparation is one in which both heart and lungs are perfused outside the body, the vessels being suitably connected to maintain a continuous circulation.

pressure or tension of carbon dioxide has been found to lie between 5 and 10 per cent of an atmosphere.

That the effect of carbon dioxide in encouraging the relaxation of the heart between beats is dependent upon the change in hydrogen-ion concentration of the perfusion fluid has been shown by securing the same results in experiments with perfusion fluids to which different quantities of weak nonvolatile acids have been added. These observations are

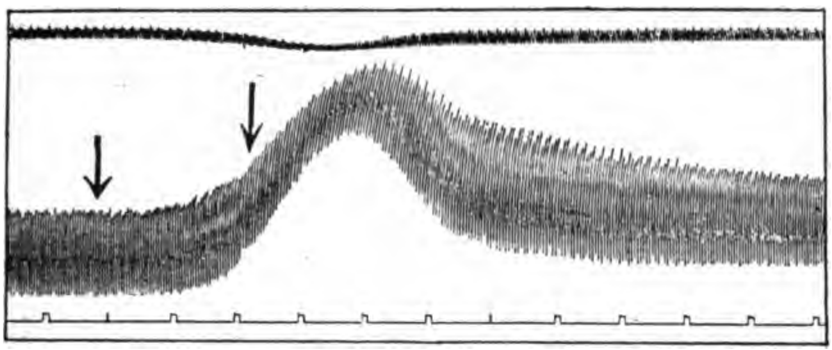


Fig. 40.—Volume curve of ventricles of cat (lower curve) in a heart-lung perfusion preparation. The air used to ventilate the lungs was replaced between the arrows by a mixture containing 20% CO_2 and 25% O_2. This caused dilatation of the ventricles along with feebler beats and a tendency for the arterial pressure to fall (upper curve). The after effect was an improvement of the beat. (From Starling.)

of practical importance because of the light which they throw on the cause of cardiac failure following upon conditions in which there has been excessive removal of carbon dioxide from the blood, as in forced ventilation of the lungs. Yandell Henderson has suggested that surgical shock may be, partly at least, due to cardiac failure following the "washing out" of carbon dioxide from the blood by the dyspnea so often incident to the administration of anesthetics in surgical operations.

CHAPTER XX

THE PHYSIOLOGY OF THE HEARTBEAT

THE ORIGIN AND PROPAGATION OF THE BEAT—THE PHYSIOLOGIC CHARACTERISTICS OF CARDIAC MUSCLE

The origin and propagation of the heartbeat are studied on the excised heart of a frog or turtle, or on the mammalian heart by perfusing it under suitable conditions, which have already been described. The results obtained on the cold-blooded heart apply more or less directly to the warm-blooded. In the first place it is clear that the rhythmic contractility of the heart is not at all dependent upon the central nervous system, for if it were so, the excised heart could not continue beating. This fact does not, however, necessarily imply that the beating power is independent of nervous structures, for in the heart itself an extended network of nerve cells and connecting nerve fibers can readily be demonstrated. It might quite well be the case that the rhythmic beat is dependent upon the transmission to the muscle fibers of the heart of impulses generated in the nerve cells and transmitted along the nerve fibers of this local nervous system. Such is the *neurogenic hypothesis of the heartbeat.*

On the other hand, it may be that these nervous structures are not at all responsible for the origination of the beat, but serve merely as stations on the pathway of the nerve impulses, transmitted to the heart from the central nervous system along the vagus and sympathetic nerves, for the purpose of altering the rate of the heartbeat so as to adjust it to the requirements of blood supply in the various parts of the body. In such a case the rhythmic power would reside in the muscular tissues of the heart—that is, each cardiac muscular cell would have the power, not merely like skeletal muscle of contracting in response to a stimulus transmitted to it, but also of originating that stimulus within itself. This is the *myogenic hypothesis.* Much controversy has raged around these two hypotheses and although space will not permit a detailed study of the question, it will be necessary, on account of the great importance of the subject from the physiologic standpoint, briefly to review the main arguments of each school of thought.

There is no piece of evidence offered by the advocates of either the neurogenic or the myogenic hypothesis that can, taken singly, be con-

sidered as absolutely conclusive. Although some of "the proofs" may at first sight appear to be conclusive, yet each of them breaks down when subjected to a closer scrutiny. It is only after we have collected all the evidence for and against each view that we shall be in a position to come to any conclusion, and even then it will be plain that our conclusion can be only tentative.

Myogenic Hypothesis

Taking first of all the evidence in support of the myogenic hypothesis, the following stands out most prominently:

1. The heart beats in the embryo chick before any nerve cells have grown into it, and not only this, but if portions of heart muscle are removed from the embryo and placed in blood plasma, they will continue beating for many days It has also been observed that cells may wander off from this mass of cardiac muscle and undergo multiplication and differentiation, so as to produce isolated muscle cells which exhibit rhythmic contractility. The rebuttal on the part of the neurogenists of this apparently unassailable evidence is to the effect that, although embryonic muscle cells may exhibit the power of rhythmic contraction, this does not mean that the fully developed muscle cells will necessarily have such power. In the eary stages of embryonic development, it is of course evident that the functions which in the fully developed animal are delegated to various special organs and tissues should be performed by cells having several such functions in common. The muscle cells of the heart, for example, may to start with be possessed of a power not only of contracting but also of initiating the contraction. It may be that they are partly nervous in character and that only later, when the differentiation is consummated, does the power of rhythmic contraction become delegated to the nervous element and that of contraction retained by the muscle itself.

2. The nervous structure in the heart may be damaged either by mechanical means or by drugs without apparently interfering with the power of rhythmic contraction; for example, in the heart of large turtles it is possible to dissect out a considerable amount of nervous tissue without any disturbance of the beat, and in all animals the administration of atropine, which paralyzes the postganglionic fibers of the autonomic nervous system (see page 226) found in the heart, does not affect it.

3. The apex of the ventricle in such hearts as that of the turtle can be shown, by careful histologic examination, to contain no nerve cells, and although a few nerve fibers may be found, these are functionless without nerve cells. This virtually nerveless piece of heart muscle can be made to contract rhythmically by perfusing it with suitable saline

solution under pressure and starting the beating by application of electrical stimuli. Isolated strips of ventricular muscle, in which also no nervous element can be demonstrated, may under favorable conditions be caused to beat quite regularly if supplied with proper nutrient fluid. The rebuttal of this evidence is twofold: In the first place, skeletal muscle itself under certain conditions, such as exposure to solutions containing an excess of phosphate (Biedermann's), may exhibit rhythmic contractility, especially on cooling, which indicates that exhibition of rhythmic power in isolated portions of cardiac muscle need not mean that under ordinary conditions such power is responsible for the normal heartbeat. In the second place, it is pointed out that although we can not reveal their presence by present-day histologic methods, this is not conclusive evidence that the heart-muscle fiber may not possess some nervous structures capable of functioning as nerve cells.

The heart even of mammals can be made to continue beating for several days after excision from the body. The nerve cells, as we know them in the central nervous system at least, can not, on the other hand, be made to functionate for more than a few hours after death. Therefore, it is argued, the heartbeat in surviving mammalian hearts can not depend on the nervous structures. The argument is however easily refuted: on the one hand, we do not know that the nerve structures situated peripherally in the heart muscles are of the same viable nature as those composing the central nervous system; and, on the other, the survival of the heart may in itself be sufficient to maintain around the nerve cells embedded in it a nutrient environment which is much more physiologic than that which we can supply in artificial perfusions of surviving nervous tissues.

4. Circumstantial but nevertheless strong evidence is furnished by the fact that many other varieties of involuntary muscle are endowed with rhythmic contractility; thus, the muscle of the intestines, of the ureters, of the bladder, of the uterus, of the blood vessels of certain animals, and of the lymph vessels in the so-called lymph hearts, maintain rhythmic contractility after isolation from the animal body. The rhythmic power seems in certain of these cases to be independent of nervous control.

Neurogenic Hypothesis

In favor of this hypothesis the following evidence is offered:

1. The heart of certain animals—of Limulus, the king-crab, for example, is definitely dependent for its rhythmic contractility upon neighboring nervous structures. The heart of this animal is a tubular sacculated organ, and along its dorsal surface there runs longitudinally a

nerve cord containing ganglion cells and giving off fibers which proceed in part directly to the heart and in part to lateral cords (Fig. 41). Removal of this median nerve cord is followed by total abolition of the heartbeat; the heart becomes perfectly quiescent like an unstimulated skeletal muscle. In appraising the evidence at its true value, it must be noted that although by stimulation of the nerve fibers contraction of the heart can be produced, the contraction is like that of a skeletal muscle—it is not rhythmic; and moreover—and this is most important—if the various physiologic properties of muscle as described below be studied (page 176), it will be found that in all of them the quiescent heart muscle behaves, not like the heart muscle of other animals, but like that of skeletal muscle. This evidence, therefore, while indisputably showing that the heart of Limulus depends for its rhythmic power upon neighboring nerve structures, does not justify the assumption that this will be the case in the heart of animals having different physiologic properties.

2. The disposition of the nervous structures in the heart, especially of the frog and turtle, exactly corresponds to the degree of development of

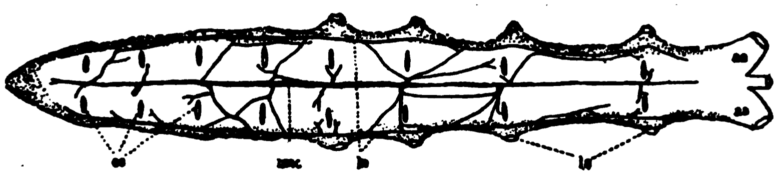


Fig. 41.—Heart and cardiac nerves of *Limulus polyphemus*. (Carlson.) *aa*, anterior arteries; *la*, lateral arteries; *ln*, lateral nerves, *mnc*, median ganglionic chain; *os*, ostii or afferent stomata, each pair of which corresponds to one of the segments into which the *Limulus* heart is divided.

the rhythmic power of the different parts of the heart; thus, the greatest rhythmic power is manifested by the sinus and the least by the tip of the ventricle at the bulbus arteriosus. In the former position the nerve structures are very prominent; in the latter, no nerve cells and but few nerve fibers can be detected. This proof is, however, easily assailed. In the first place, it may merely be a coincidence that the disposition of the nerve structures and the development of rhythmic power correspond. The unequal rhythmic powers may depend primarily on a difference in structure of the muscle fibers themselves, such differences having been shown to exist between the muscle cells of the sinus and those of, say, the ventricle. The former cells, for example, have much less developed crossed striation and their protoplasm is much more granular; in short, they are much more embryonic in type than the cells from the tip of the ventricle.

If a jury had to return a verdict from evidence of so conflicting a character, it would no doubt be equivalent to that of the Scottish court—"not

proven." But it is likely that the majority of the jury would vote in favor of the myogenic hypothesis. Probably the safest viewpoint to take at the present time is that the power of rhythmic contraction is inherent in the cardiac muscle fibers, being most highly developed in those of the venous end of the heart, and least developed in those of the arterial end. Such a conclusion does not deny to the nervous structures of the heart the power under certain conditions of also assuming rhythmic activity. In one case at least—namely, the heart of Limulus—we know that this is so. For some reason in this animal the cardiac muscle fiber has lost its inherent rhythmic power, and is now dependent for its activities upon rhythmic nervous discharges transmitted to it from the neighboring nerve cords, a condition which is paralleled in the higher animals in the innervation of the respiratory muscles. The respiratory center rhythmically discharges impulses to the muscles, which are quiescent in the absence of these impulses.

The Pacemaker of the Heart and Heart-block

In a volume of this nature, devoted primarily to the practical application of physiology, the discussion of these problems may seem a little out of place, but that this is not the case is seen when we consider that the experiments upon which the various points of evidence depend bring to light facts of the very greatest importance in the study of the physiology of the heartbeat. One fact which stands out prominently is that *the greatest rhythmic power resides in the basal portion of the heart*—that is, in what corresponds, in the more primitive hearts, to the sinus venosus.

Although the muscle of the entire heart possesses rhythmic power, it does not do so to an equal degree; in the sinus the rhythmic power is extraordinarily developed, while in the bulbus arteriosus it is scarcely recognizable. This observation suggests the possibility that the sinus may dominate the heartbeat—that it may be the "pacemaker" for the heart as a whole. The most natural method for demonstrating such a possibility would be to observe the effect on the heartbeat of *some block* between the sinus and the rest of the heart. Such a block can be introduced in the heart of cold-blooded animals by local compression around the various junctions. If a thread is tied around the sinoauricular junction, the sinus will go on beating uninterruptedly, but the auricles and ventricles—that is, the greater part of the heart below the ligatures—will cease beating, sometimes entirely (Stannius' ligature). After a while, however, the heart below the ligature will usually begin to beat, but at a rhythm which is slower than, and independent of, that of the sinus.

The experiment can be still better performed by using a wedge-shaped clamp. (Gaskell's clamp.) If this is applied so that the heart can be pinched either at the sinoauricular junction or at the auriculo-ventricular, it will be found that, as the cardiac tissue is gradually

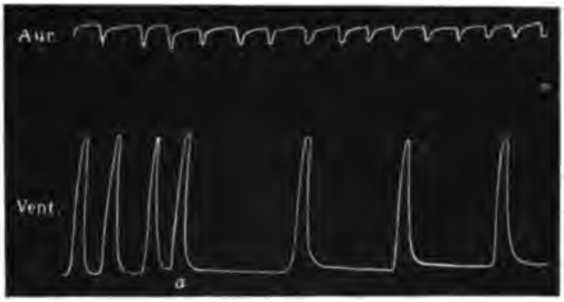
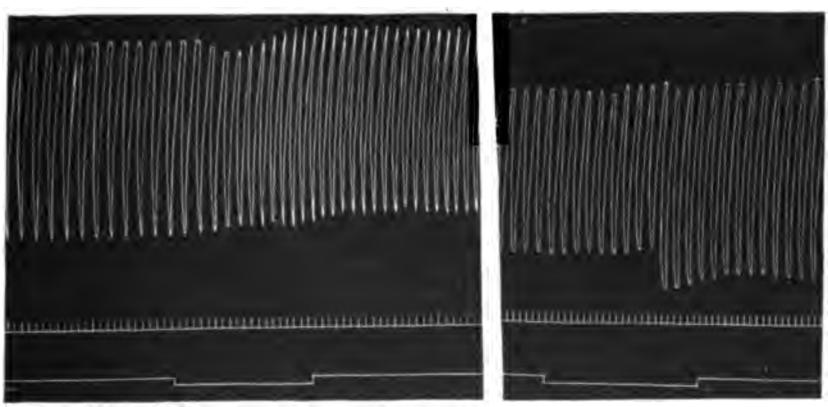


Fig. 42.—Heart-block produced by applying clamp at a-v junction. The clamp was tightened at *a*. (From Brubaker.)

pinched, the portion of the heart below fails to beat as quickly as that above the clamp (Fig. 42). This is known as *partial heart-block,* and the degree of the block is indicated by the numerical expression 2 to 1, 3 to 1, 4 to 1, etc., meaning that the sinus is beating either twice as quickly as the ventricle, or three times, or four times as the case may



1. *2.*

Fig. 43.—Tracing of contraction of ventricle, showing the effect of the local application of heat to the auricle at *1*, and to the apex of the ventricle at *2*. Note that the rate increased in the former case.

be. Similar conditions of heart-block may also be produced by cutting the cardiac tissue partly across at various places in the heart.

Further evidence that the sinus dominates the beat in the heart of

cold-blooded animals is furnished by observing the effects of local heating or cooling of the various parts of the heart. In all rhythmically acting structures it is well-known that heat increases the rate of the rhythm and cold depresses it. If we locally warm the region of the sinus, as by holding a heated wire near it the whole heart will immediately beat quicker; but if we locally heat the tip of the ventricle, no alteration of rhythm will be observed to occur (Fig. 43).

The establishment of the fact that the sinus dominates the heartbeat —that it is the pacemaker of the beat—raises the question as to how the impulse originated at this place is transmitted over the rest of the heart, and here again a neurogenic and a myogenic hypothesis have to be considered. Before going into this question, however, it will be well for us to consider briefly the manner of response of cardiac muscle fiber to a stimulus, because the behavior of cardiac muscle under such conditions is considerably different in many regards from that of skeletal muscle, and it is to these differences that many of the peculiar alterations in the beat observed after interfering with the conducting structures between the sinus and the rest of the heart, are to be explained.

The Physiologic Characteristics of Cardiac Muscle

It is necessary to bring the heart into a quiescent state in order to investigate the properties of its musculature. This is accomplished by the application of the Stannius ligature between the sinus and the auri-

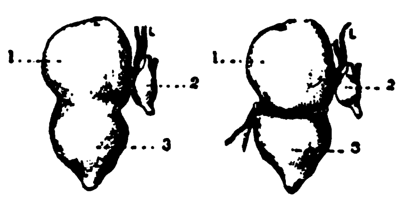
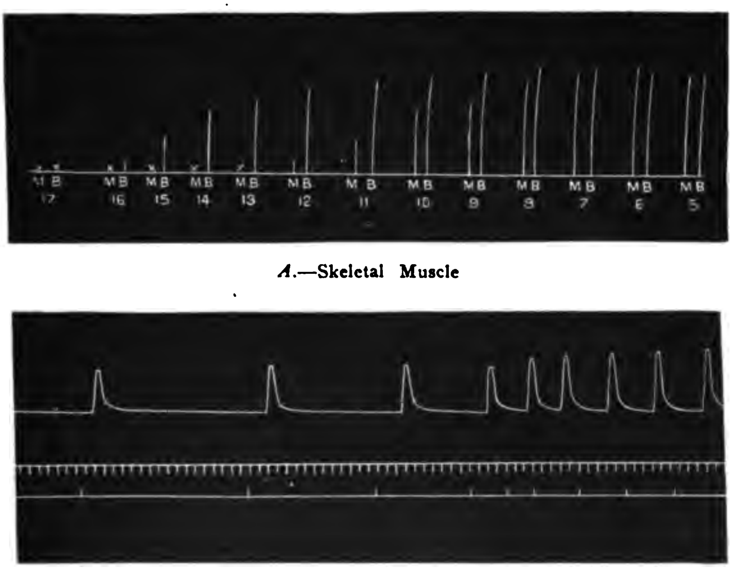


Fig. 44.—Frog heart showing the position of the first and second ligatures of Stannius (Hedon): *1*, auricles; *2*, sinus; *3*, ventricle. It is the first ligature which brings the heart to standstill.

cles (Fig. 44). After tightening the ligature the auricles and ventricles become quiescent, and by observing the effects produced by the application of electric or other stimuli we can compare the behavior of the cardiac muscle with that of skeletal muscle similarly stimulated. This comparison is made because of the assistance which it offers in comprehending the properties of cardiac muscle. As a matter of fact, recent investigations have shown that the differences between the two types of muscle are not fundamental, since under certain conditions the one may

be made to behave like the other. They are dependent upon the presence or absence of anastomosis between the muscle fibers.

1. When electric stimuli of varying strengths are applied to skeletal muscle, the contraction produced by each stimulus is proportional to the strength of the latter until this has become of such a strength that the maximal response is elicited. In cardiac muscle, on the other hand, an entirely different result is obtained, for the weakest stimulus, if it produces any response at all, produces one that is maximal; that is, the height of contraction is the same as it would have been had a much stronger stimulus been applied. Expressing this result in general terms, we may say that in cardiac muscle *a minimal stimulus produces a maximal effect,* whereas in skeletal, the effect, as measured by the height of contraction, is proportional to the intensity of stimulation. This is sometimes known as the "all or nothing phenomenon" (Fig. 45).



A.—Skeletal Muscle

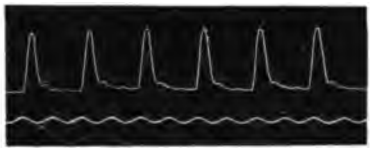
B.—Cardiac Muscle

Fig. 45.—Effects of stimuli of increasing strength on skeletal and cardiac muscle to illustrate the "all or nothing" principle in the latter. (From *Practical Physiology*.)

2. If maximal stimuli are applied successively and at short intervals of time to skeletal muscle, a slightly higher response results from each succeeding stimulus, until about ten stimuli have been applied, after which for some considerable time the same height of contraction follows each stimulus. If each contraction is recorded, it will be seen that the first few contractions give a staircase effect; that is, if a horizontal line is drawn from the top of each contraction to the next one, the effect of a

staircase with gradually diminishing steps will be produced. If we repeat this observation with cardiac muscle, we shall find that the staircase phenomenon or *treppe,* as it is called, is very pronounced; and moreover, in obedience to the all or nothing principle, the treppe is obtained in cardiac muscle whatever may be the relative strengths of the stimuli applied to the heart, provided always that all of them are effective; whereas in the case of skeletal muscle it can be demonstrated only provided the stimuli are of equal strength (Fig. 46).

3. If an effective stimulus is applied to a skeletal muscle while in process



Skeletal muscle

Cardiac muscle

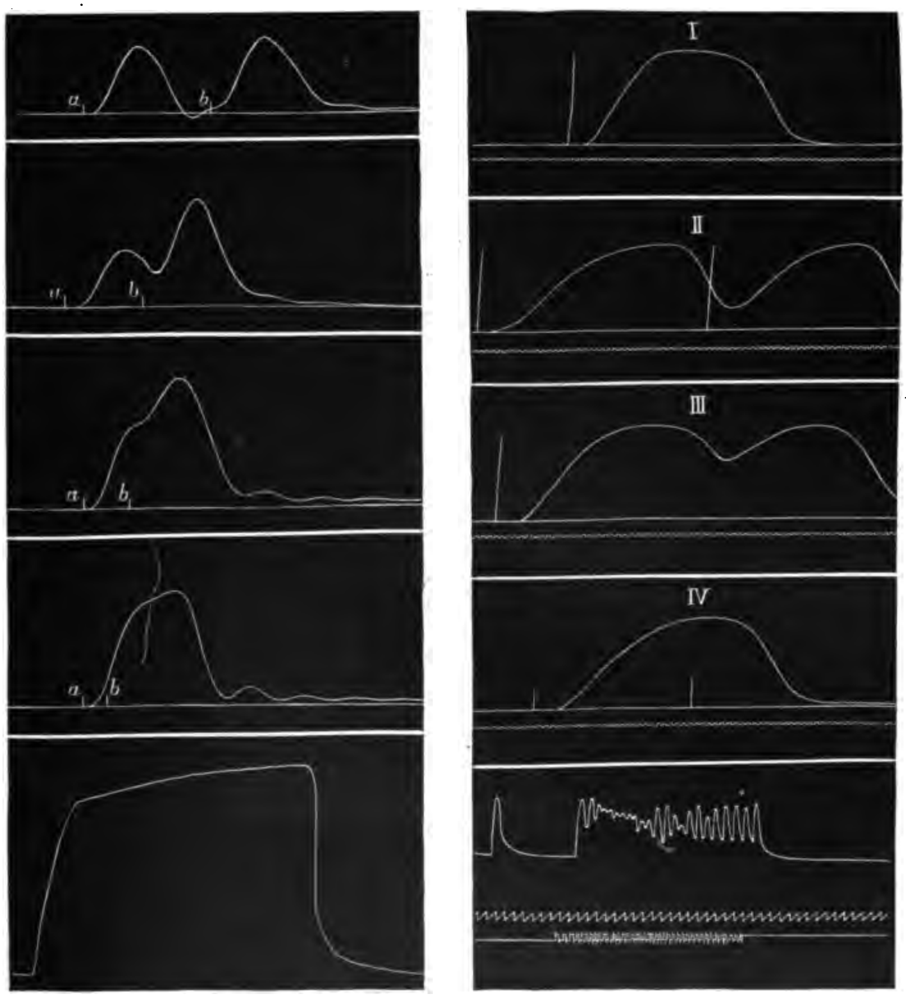
Fig. 46.—The effects of successive stimuli on skeletal and cardiac muscle to show the prominence of the staircase phenomenon, or *treppe,* in the latter. (From T. G. Brodie.)

of contraction, as in response to a preceding stimulus, the second stimulus prolongs the contraction produced by the first one. If, however, the second stimulus is applied during the latent period* of the first one, it will have no effect—that is, the muscle during this period is refractory.† From these results it follows that stimuli succeeding each other during the contraction period will, in the case of *skeletal* muscle, cause a continuous contraction, or tetanus, as it is called, because the contraction produced by each stimulus will add itself to that of its predecessor before any trace of relaxation has set in. If, however, the second stimulus is applied so late in the contraction period of the first that time is not available for the latent

*By "latent period" is meant the period after the moment of application of a stimulus during which no effect of that stimulus is observed.

†By "refractory period" is meant the time following the application of a stimulus during which a second stimulus develops less than its full effect or no effect at all.

period to expend itself, then obviously a slight relaxation will have occurred before the effect of the second stimulus develops itself, and tetanus will be incomplete. These facts will be evident from the accompanying tracings (Fig. 47).



Skeletal muscle

Stannius' heart

Fig. 47.—The effects of successive stimuii and of tetanizing stimuli on skeletal muscle and cardiac muscle. The small vertical marks show when the stimuli were introduced. (Compiled from tracings published by T. G. Brodie and Leonard Hill.)

In the case of *cardiac* muscle the above described properties are quite different, for *the refractory phase extends throughout the whole period of contraction;* that is, a second stimulus applied during the contraction produced by a previous stimulus has no effect whatsoever; it does not

have one until the muscle has reached the full extent of its contraction and is about to relax. Since a latent period must supervene upon the application of this second stimulus, it follows that no complete fusion of the contractions is possible. Complete tetanus therefore, does not occur in cardiac muscle, however frequently the stimuli may be applied (Fig. 47).

The refractory phase is a property of extreme importance in understanding many of the peculiar irregularities observed in cardiac action. If we observe the effect of stimuli applied at varying periods after the

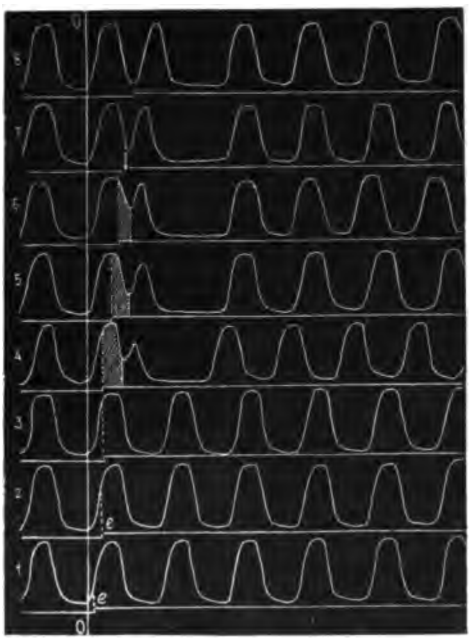


Fig. 48.—Myograms of frog's ventricle, showing effect of excitation by break induction shocks at various moments of the cardiac cycle. The line *O* indicates the commencement of all the beats during which the shock is sent in. It will be noted that in *1*, *2* and *3*, the heart is refractory to the stimulus. The signals indicate the moments at which the stimuli were applied. From *4* to *8* the heart reacts by an extrasystole, after a delay, which is progressively less the later in diastole the stimulus enters, as shown by the sections shaded obliquely to make them more conspicuous. The extrasystoles increase in height from *4* to *8*, each being followed by a compensatory pause. (From Luciani's *Human Physiology*.)

termination of the refractory phase of a previous stimulus, we shall find that the height of the extra contraction is directly proportional to the time after the end of the refractory period at which it is applied. If a stimulus is applied at the very beginning of diastole, the extra contraction will be small, whereas if it is applied at the end of diastole, the extra contraction will be at least as high as that of the preceding. It may be higher because of the treppe.

These observations enable us to interpret the results obtained by *applying electric shocks* (*extra stimuli*) *to the beating heart* during different phases of systole and diastole. During systole, the muscle being refractory, no effect is produced by the extra stimulus, but during diastole extra systoles which are progressively more pronounced the later in diastole they occur, follow the application of each stimulus. These results are so far exactly like those obtained with a quiescent heart. But another phenomenon now becomes evident; namely, that following each extra systole there is a compensatory pause in the action of the heart, of such duration that, when the next natural beat occurs, it does so practically at the same time as it would have occurred had no artificial stimulus been applied. This will be apparent from the following diagram (Fig. 48).

It should be noted that the refractory period is greatly diminished by raising the temperature of the heart. Indeed, under these conditions and with strong stimulation it may be possible to produce an almost complete tetanus.

The importance of knowing the above facts is that we are thereby enabled to explain the peculiar manner in which the ventricle responds to stimuli transmitted to it from the sinus and the auricle. The musculature of the auricle and ventricle of the mammalian heart is not one continuous sheet, but is separated by a space at the auriculoventricular junction, across which, in specially organized structures, the beat of the auricle is transmitted to the ventricle. Sometimes the stimuli are so frequent that the ventricular muscle is unable to respond to each stimulus transmitted to it, with the result that marked irregularities in contraction occur (see page 280). In this way certain of the cardiac irregularities observed in man can be explained. Thus, the so-called *pulsus bigeminus* is due to every second beat being an extra systole. This second beat is therefore generally weaker than the preceding and succeeding normal beats, and it is almost always followed by a compensatory pause. When the intervals separating the beats are of uniform length, although every second beat is diminished in size, the pulse is termed *pulsus alternans.*

CHAPTER XXI

THE PHYSIOLOGY OF THE HEARTBEAT (Cont'd)

THE ORIGIN AND PROPAGATION OF THE BEAT IN THE MAMMALIAN HEART

As has been shown in the preceding chapter, there is no doubt that in the *cold-blooded heart* the beat originates at the sinus venosus, whence it spreads to the rest of the heart. Very strong evidence has also been presented to indicate that the beating power is inherent in the muscle fiber itself and independent of nervous structure. This would suggest the further possibility that the structures through which the beat is propagated are the muscle fibers and not the nerve fibers—in other words, that the propagation of the heartbeat, like its origination, is myogenic rather than neurogenic. Direct proof of this hypothesis is readily furnished by numerous experiments, among which may be mentioned making interdigitating cuts across the heart, or excising a ribbon of ventricular muscle by an incision simulating the walls of Troy. In both these cases the beat will be found to travel from one end of the muscular band to the other, although it is evident that all the nerves proceeding from base to apex of the heart must have been severed. Of course this evidence is not irrefutable, for it might be argued that there are nervous structures disposed in the form of a plexus continuously all over the heart, and that some branches of the plexus remain uncut in the above experiments. It is only in the heart of Limulus that undoubted evidence exists that the beat is transmitted by nerves, but as we have seen, this heart in all its properties is probably the proverbial exception which proves the rule. The balance of evidence stands in favor of the view that the propagation of the beat over the cold-blooded heart is myogenic and not neurogenic.

CONDUCTING TISSUE IN MAMMALIAN HEART

When we attempt to investigate the problems of the origin and propagation of the beat in the *warm-blooded heart*, many experimental difficulties of course face us. In overcoming these, the first thing we must do is to establish the structural relationship between cold-blooded and warm-blooded hearts. In the embryo of both classes of animals the

heart arises as the so-called cardiac tube. As development proceeds, diverticula grow out from the walls of this tube to form the auricles and ventricles. In the comparatively simple heart of the turtle these dispositions of the auricles and ventricles in relationship to the cardiac tube are more or less evident even in the fully developed heart, particularly in the case of the auricles (Fig. 49); but in the heart of the higher mammalia it is impossible by superficial examination alone to show any remains of the primitive cardiac tube. More careful anatomic investigations during recent years have, however, shown that it exists in the form of certain definite structures composed of tissue histologically quite different from that of the rest of the heart, and disposed in such a manner

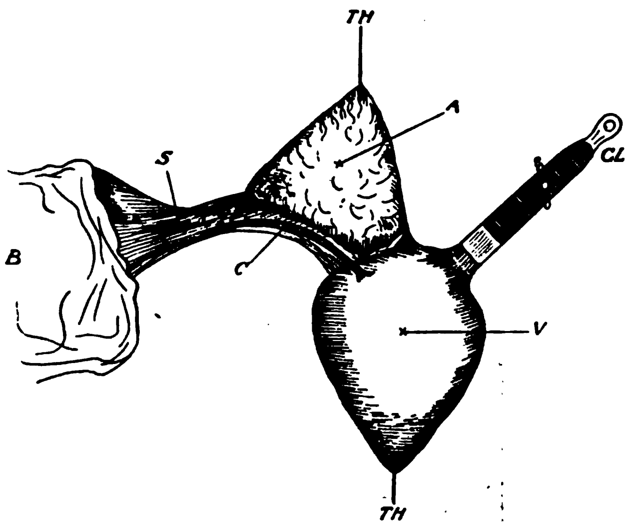


Fig. 49.—Heart of tortoise as suspended. *B*, body of tortoise; *TH*, threads to levers; *CL*, clamp holding aorta; *A*, auricle; *C*, coronary nerve; *S*, sinus; *V*, ventricle. (From Gaskell.)

as would indicate not only that it is derived from the primitive cardiac tube, but also that it is the main pathway along which the beat is transmitted.

This primitive cardiac tissue is much better developed in certain regions than in others, the first portion of it to be discovered being that known as the *auriculoventricular node*, or the node of Stanley Kent* (Figs. 50 and 51). This structure is found at the base of the interauricular septum on the right side and near its posterior margin. It exists as a collection of peculiar small primitive cells and fibers, and is continued downward as *a bundle* of the same peculiar tissue to the interventricular septum, where, near the union of the posterior and median flaps of the aortic

*The discovery of this node is often erroneously attributed to His, and called after his name.

valve, it bifurcates so as to send a branch down each side of the septum immediately below the endocardium. Each main branch, as it proceeds downward on the septum, divides up into an intricate system of smaller branches, which become reflected over the inner surface of the ventricles, where their existence has been known for some time as the so-called

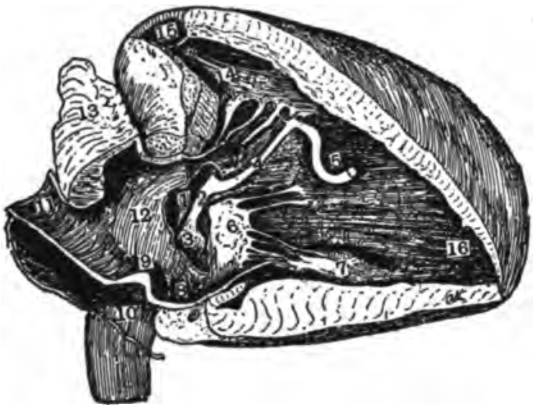


Fig. 50.—Dissection of heart to show auriculoventricular bundle (Keith); 3, the beginning of the bundle, known as the A-V node; 2, the bundle dividing into two branches; 4, the branch running on the right side of the interventricular septum. (From Howell's *Physiology*.)

Fig. 51.—Photograph of model of the auriculoventricular bundle and its ramifications, constructed from dissections of the heart (Miss De Witt). All of the branches in the left ventricle are not included. (From Howell.)

Purkinje fibers. The fibers ultimately end in close association with the papillary muscles. The node and main bundle and the two branches before they have begun to divide are surrounded by fibrous tissue, and they seem to have a liberal blood supply. It is of interest that they contain a high percentage of glycogen. In the human heart the auriculo-

ventricular node and bundle measure about 15 mm. in length and about 2 mm. in width.

The rest of the tissue between the auricles and ventricles is fibrous in nature, although other connections like those of the auriculoventricular bundle have been described by Kent. One of these, called *the right lateral connection,* runs between the right auricle and the external wall of the right ventricle.

Another, but much smaller, mass of similar embryonic cardiac tissue has more recently been discovered by Keith and Flack in the parts of the auricle which correspond anatomically to the sinus venosus of the heart of cold-blooded animals—that is, in the area lying between the openings of the venæ cavæ and around the coronary sinus. To be more explicit, this tissue lies "in the sulcus terminalis just below the fork formed by the junction of the upper surface of the auricular appendix with the superior vena cava." This *sinoauricular node,* as it is called, is more or less club-shaped, the blunt end of the club being above, as shown in the accompanying figure (Fig. 52). It is important to note that there is no direct connection visible between the sinoauricular and auriculoventricular nodes (Fig. 53).

Another anatomic fact seen also in the accompanying figure, concerns the disposition of the muscular fibers of the auricle. These radiate in bundles in a peculiar fan-shaped manner from a point which lies immediately below the sinoauricular node to all parts of the superficies of the right auricle. This point has been called the *concentration point.* At the termination of the venæ cavæ, the muscle fibers are arranged more or less circularly.

Having become familiar with the disposition in the mammalian heart of the primitive cardiac tissue, along which in the heart of the lower animals we know that the heartbeat spreads, we may now proceed to examine the evidence showing that this tissue is also responsible for the origination and propagation of the beat in the heart of mammals. With regard to the origin of the beat in a normally beating mammalian heart, it is of course impossible to tell where this takes place. If the heart is excised, however, it will continue to beat for a few moments, and as it dies it will be observed that the power of contraction remains in the auricular region, and *particularly* at the bases of the venæ cavæ, for a considerable time after the ventricles have ceased to beat. This part—the *ultimum moriens*—is situated in most hearts somewhat lower than the sinoauricular node. That it is the last part of the heart to cease contracting does not necessarily mean that it is the part of the heart in which the beat ordinarily originates; it means simply that this is the part of the auricle in which the power of contraction remains for the

longest time after death. Although the observation does not enable us to determine exactly where the heartbeat originates, yet it makes it very probable that this is somewhere in the auricles; a conclusion which is borne out by many other pieces of evidence, such as those obtained by

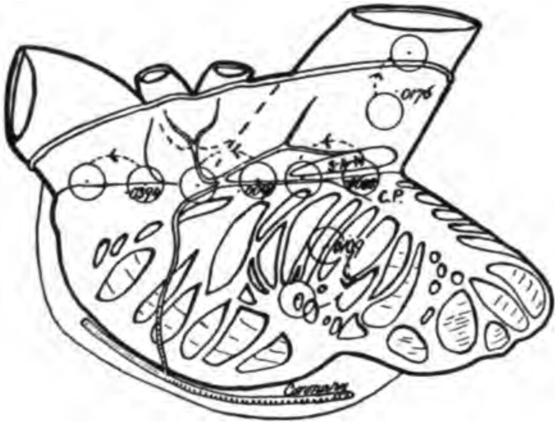


Fig. 52.—Diagram of an auricle showing the arrangement of the muscle bands; the concentration point (*C.P.*); and the outline of the *S.A.* node (*S.A.N.*). The diagram is to scale, and illustrates by the circles and connecting dotted lines the method of leading off by paired contacts and the subsequent orientation. (From Thomas Lewis.)

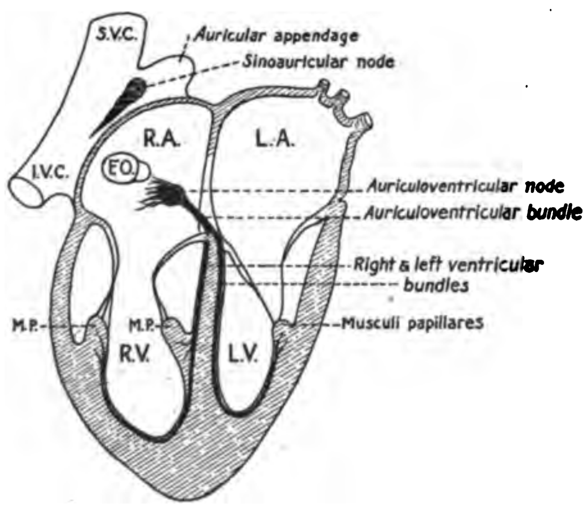


Fig. 53.—Diagram to show the general ramifications of the conducting tissue in the heart of the mammal. It will be observed that there is none of this tissue between the sinoauriculo- and auriculoventricular nodes.

the study of polysphygmograms (page 273), of electrocardiograms (page 266), and of observations on the heart during heart-block (page 270). Our problem therefore narrows itself down to determining the exact point of the right auricle at which the beat originates.

SITE OF ORIGIN OF THE BEAT

The working hypothesis from which we may proceed to attack this problem is that the beat originates in the sinoauricular node, and to put this to the test, various methods have been employed: (1) Warming or cooling or injuring the node and noting the effect on the heartbeat. Such procedures greatly affect the rate of the heartbeat, whereas they produce no change when applied to other parts of the heart. (2) Determination of the comparative rhythmic power of strips cut out from different regions of the auricular walls. It is greatest in those taken from the region of the node. (3) Determination by the use of galvanometric curves of the relation of the node to the seat of origin of cardiac impulse. By all these methods the results indicate clearly that the beat originates in the sinoauricular node, but on account of the great importance in connection with the interpretation of electrocardiograms in man, it is particularly with the result of the third group of experiments that we will concern ourselves here.

Evidence Furnished by Studying the Current of Action Which Accompanies the Heartbeat

To start with, it is essential that we make ourselves familiar with *the principles of the methods* employed. These principles are briefly as follows: When a wave of contraction passes along a muscle, it is immediately preceded by a change in electrical potential, which can be detected by means of a galvanometer connected with the muscle through so-called nonpolarizable electrodes. The galvanometer employed must be extremely sensitive, and must not vibrate after the current has ceased to pass. The form generally in use today is known as *the string galvanometer of Einthoven*. It differs from the galvanometer ordinarily employed in physical laboratories in that the current instead of passing through a coil of wire surrounding a magnetic needle, passes through a silverized quartz thread suspended in the strong magnetic field which exists between the two opposing poles of a horseshoe electromagnet. The string is thus surrounded on all sides by innumerable lines of force extending between the two poles of the magnet. When a current, however small, passes along the string, it will generate lines of force of its own, and these by reacting with the stationary lines of force of the field will cause the string to move. The string is placed in the pathway of a strong beam of light, and its shadow, after being magnified by lenses, is projected on a moving photographic plate or paper arranged in a suitable holder. The nonpolarizable electrodes referred to are employed in place of ordinary electrodes in order to obviate the generation of elec-

tric currents set up by the contact of metal with the saline constituents of the muscle juices.

If we connect a galvanometer by means of nonpolarizable electrodes with two parts of a denervated muscle (the curarized sartorius of the frog), it will be found that a current is set up whenever a wave of contraction passes over the muscle from one end to the other. The part which first contracts becomes electrically negative to the rest of the muscle, but as the wave of contraction passes along it, the "negativity" decreases at the end at which the wave started until, when the wave has reached the middle of the strip, neither end of the muscle shows any difference in potential, so that the string comes back to a position of rest. However, as the contraction wave reaches the farther end of the muscle, this lead in turn becomes negative, and the string swings in the

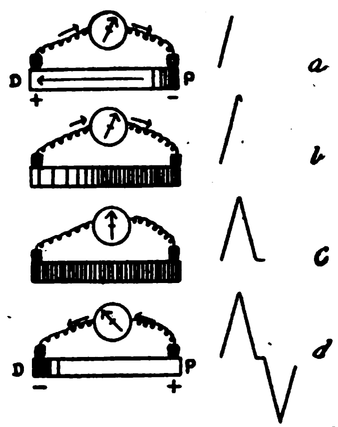


Fig. 54.—Diagram to illustrate the development and spread of the wave of negativity in a strip of muscle (curarized sartorius) when stimulated at the end (*P*). The shaded portions show the position of the negativity. The portion of the curve drawn by the deflections of the galvanometer at each stage are shown at the right (*a*, *b*, *c*, and *d*). (After Lewis.)

opposite direction (Fig. 54). From this comparatively simple experiment it can be seen that a muscular contraction wave arises at the electrode which is negative first, and that the movement of the string of the galvanometer is most marked—that is, the deflection is greatest—when the two electrodes are applied at the extreme ends of the muscle. When they are brought closer together, the initial deflection becomes much less marked; in other words, the amplitude of the negative wave is greatest when the time interval between the receipt of the excitation at the two contacts is greatest.

The application of these facts to the study of the initiation of the beat in the auricle requires that we should consider another proposition: namely, if a pair of contacts are arranged in the center of a circular sheet of muscle and the edge of this sheet is stimulated at different

points, the amplitude of deflection of a galvanometer connected with the pair of contacts will be most pronounced when these are radial to the points of stimulation, for under these conditions it is evident that the greatest possible difference will exist between the intervals required for the wave to reach each contact.

Bearing these principles in mind, we may now proceed to examine the *evidence pointing to the origin of the heartbeat at the sinoauricular node:* (1) When two electrodes are applied at different points of the auricle, the amplitude of movement of the string of the galvanometer produced by each heartbeat is greatest when the line joining the electrodes converges on the sinoauricular node. To make this clear the movement of the string must be photographed in the manner above described, the resulting tracing being called an electrocardiogram. From the experiments with the circular sheet of muscle alluded to it is evident that the stimulus to produce this result must have arisen in the neighborhood of the node. (2) If one electrode is placed on the sinoauricular node and the other electrode is moved about from place to place on the auricle, the deflection being noted at each new position, the electrode on the node will always be found to be negative to the other electrode;* which, however, will not be the case if both electrodes are moved about on other parts of the auricle.

(3) As we shall see immediately, the current of action of the beating heart may be recorded by connecting a galvanometer with various parts of the body; for example with the right fore limb and the left hind limb. On the electrocardiogram thus obtained are several waves, one of which, called the P-wave, can easily be shown to correspond to the contraction of the auricle (see Fig. 261). If now we compare such electrocardiograms with those obtained while contractions of the auricle are produced by applying artificial stimulation to various parts of it, it will be found that the artificial simulates the normal curve only when the stimulated part is in the neighborhood of the sinoauricular node. In other words, it is only when the stimulus is applied to the sinoauricular node that a characteristic P-wave is obtained. When the appendix or the superior vena cava is stimulated, the P-wave is distorted although the other waves of the electrocardiogram may be normal.

(4) By taking simultaneous electrocardiograms from direct leads placed on the auricle and comparing the record with that of a standard limb lead taken simultaneously, we shall find by exact measurement that the time of onset of the excitation wave of the auricle, as measured in relationship to the P-wave on the standard electrocardiogram, is shortest

*The connections between the electrodes and galvanometer are always arranged so that any upward movement of the shadow of the string above the line of equal potential at the two electrodes indicates electric negativity.

when one electrode is over the upper end of the sinoauricular node, and that in other regions of the auricle it always appears at a later interval. Further details on this subject will be found in the papers by Eyster and Meek[8] and in Lewis, monographs.

Frequently, in taking electrocardiograms from different parts of the auricle, it is found that certain of the curves show small waves of positivity below the line of equal potential preceding the main wave of negativity. These initial deflections are most marked when both the electrodes are far removed from the sinoauricular node—for example, when they are placed on the auricular appendix; but they are never present when

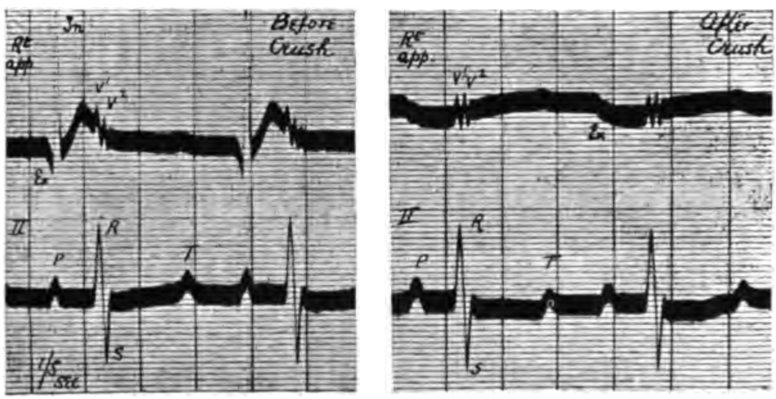


Fig. 55.—Simultaneous electrocardiograms to show the cause for extrinsic deflections. The upper curves are from the appendix and the lower ones from lead II. The chief or intrinsic deflection (*Tn*) is seen to disappear in the right-hand appendix electrocardiogram, because the base of the appendix has been crushed. The extrinsic deflection (*Ex*) remains, as do the ventricular deflections (V^1 V^2). (From Lewis.)

one of the electrodes is placed on the sinoauricular node itself. In other words, curves taken from leads at a distance from the sinoauricular node are more or less composite in form, being made up partly of the main deflection due to the arrival of the excitation and partly of the secondary deflections dependent upon extrinsic influences acting on the electrodes; that is, the electrode picks up electric discharges from distant areas of muscle while these are in a condition of contraction (Fig. 55). From these considerations it follows that the intervals between the intrinsic and extrinsic deflections should be longest in leads that are farthest from the node, and gradually become less as one of the contacts approaches the node, until over this structure the extrinsic deflection is no longer recorded. Such has been found to be the case.

CHAPTER XXII

THE PHYSIOLOGY OF THE HEARTBEAT (Cont'd)

THE ORIGIN AND PROPAGATION OF THE BEAT (Cont'd)—FIBRILLATION

Mode of Propagation in the Auricles

From the mass of evidence we have little doubt that the heartbeat originates in the sinoauricular node, and the question now presents itself as to how the beat is propagated over the remainder of the auricles and into the ventricles. Regarding the propagation of the beat over the auricles, two possibilities exist: (1) it may spread uniformly over the muscular tissue of the auricular wall until it reaches the auriculoventricular node, or (2) there may be laid down between the sinoauricular and the auriculoventricular node a special strand of highly conducting tissue. It is no argument against this second possibility that we should so far have been unable by histologic methods to differentiate any such structures.

There is considerable practical importance attached to the solution of these questions, particularly with regard to the cause of certain types of cardiac arrhythmia, such, for example, as that known as *nodal rhythm*. Thus, it is evident that if the beat is transmitted uniformly over the muscular tissue of the auricle, then the whole auricle will have contracted before the beat has reached the auriculoventricular bundle, by which it is then transmitted to the ventricles. On the other hand, if the beat should travel between the two nodes by special conducting tissue, then the impulse will have arrived at the auriculoventricular node before the auricle has contracted. As a matter of fact, it is not quite settled yet as to which of these two views is the correct one, although the balance of evidence seems to favor the former—that is, that the wave is transmitted uniformly over the muscular tissue of the auricle. (Lewis.)

The methods employed in attacking the problem have been essentially the same as those described above. One of them may be called the direct, the other the indirect. In the former, a series of pairs of contacts is placed on the auricle, each pair being in a radial direction to the sinoauricular node. The time at which the excitatory process arrives at that contact of each pair which is proximal to the sinoauricular node is accu-

rately determined from the galvanometric record. The exact distance between the contact and the sinoauricular node is then measured and from the data the average transmission time is estimated. From his results Lewis[3] concludes that the transmission rates are uniform from the node to all parts of the auricle, with the exception of the superior vena cava, in which the rate is considerably lower. One thousand millimeters per second represents very fairly the average rate at which the excitation wave travels. On the other hand, Eyster and Meek[8] state that the wave is propagated throughout the sinus node, and that it spreads to the contiguous venæ cavæ and to the auriculoventricular node with considerable rapidity, reaching the mouth of the superior vena cava in 0.01 second, whereas its passage to the auricle itself takes 0.02 second. There is therefore a delay in the passage of the wave to the auricle, which indicates that the excitation must spread to the auriculoventricular node before involving the right atrium. These authors conclude that "this leads to the inevitable conclusion that the cardiac impulse spreads to the ventricle and to the right auricle by different paths, and does not pass to the ventricle through the auricle, as ordinarily stated."

In the second, or indirect, method, the onset of the negative wave from different leads in the auricle is compared against a standard. For the standard Eyster and Meek have used the record of the mechanical systole of the auricle, but the interpretation of the result is extremely difficult on account of the rate at which the changes are occurring. Lewis, on the other hand, has used the standard electrocardiogram for purposes of comparison.

Mode of Propagation of the Beat to the Ventricles

After reaching the auriculoventricular node, the beat is transmitted to the ventricles along the auriculoventricular bundle—a fact which has been most clearly demonstrated by the experiments on *heart-block*. We have already seen (page 174) that although each chamber of the heart of a turtle or frog has a rhythm of its own, this is much more pronounced at the venous end of the heart, and when the transmission of the beat to the ventricles from the auricles is obstructed or blocked, as by compression or partial cutting at the auriculoventricular junction, the ventricles, after coming to a standstill for a time, subsequently contract with a rhythm which is entirely independent of that of the auricles.

In the mammalian heart the same results may be obtained by arranging a clamp so that it compresses practically nothing but the auriculoventricular bundle (Erlanger.) If the compression is extreme, the rhythm of the ventricles is quite independent of that of the auricles, but if it is only partial, the ventricular systoles follow regularly every sec-

ond, third, or fourth auricular contraction. If after such a complete or partial heart-block has been instituted, the clamp is removed, it will usually be found that the heart-block disappears and the auricular and ventricular contractions fall back into their usual sequence. The importance of this discovery, apart from its physiologic interest, rests in the fact that it is *exactly duplicated in clinical experience.* If the pulse tracing of the radial artery is compared with that of the jugular vein in certain types of heart disease, it will be found that the auricle is beating two or three times more quickly than the ventricles. In many of these cases it has been found on autopsy that definite lesions often syphilitic in nature involve the auriculoventricular bundle. In other cases, however, such lesions have not been discovered. Sometimes the bundle is so severely diseased that the block is complete, the ventricles contracting quite independently of the auricle (Stokes-Adams syndrome.) In such cases it is assumed that the beat originates in the uninjured part of the bundle below the seat of the block. It should be pointed out here, however, that all cases of slow pulse in the arteries are not necessarily dependent upon heart-block, but may depend upon a slow beat of the auricle itself. This is called *bradycardia.*

Sometimes after complete destruction of the auriculoventricular bundle the beat continues to be transmitted to the ventricle, and conversely this transmission has sometimes been observed to be upset by lesions not affecting the bundle. The explanation of both of these exceptional results almost certainly is that the right lateral connection described above (page 184) is serving as the main pathway of transmission for the beat.

The facility of conduction through the auriculoventricular bundle is subject to alteration by the impulses passing to it along the vagus nerve, particularly the left vagus. It can also be altered by certain drugs, especially digitalis and strophanthin. The clear demonstration that it is along this bundle that the beat is transmitted is strong evidence in favor of the myogenic hypothesis (page 171) concerning the transmission of the heartbeat, but it does not necessarily disprove the neurogenic hypothesis, for histologic investigation has shown that the bundle is closely surrounded by an intimate plexus of nerve fibers.

Spread of the Beat in the Ventricle

After the impulse has been transmitted by the bundle into the ventricles, it spreads along the many branches into which, as we have seen, the two main divisions of this bundle separate. The first part of the ventricular musculature to contract is therefore located near the termination of these branches, at the papillary muscles. That these should contract before the rest of the muscle of the ventricles, has an obvious

significance in connection with their function of tightening the chordæ tendineæ so as to prevent any bulging of the flaps of the auriculoventricular valve into the auricles when, at the beginning of the presphygmic period, the high intraventricular pressure is brought to bear on their under surfaces. After starting at this point in the ventricle, the contraction wave seems to spread farther through the ventricular muscle at a fairly uniform rate.

Investigation of this problem by means of the galvanometer has been technically a very difficult matter, and the details of the researches by Lewis and his pupils have not as yet been published in full. According to the preliminary communications at hand, however,[8a] it appears that,

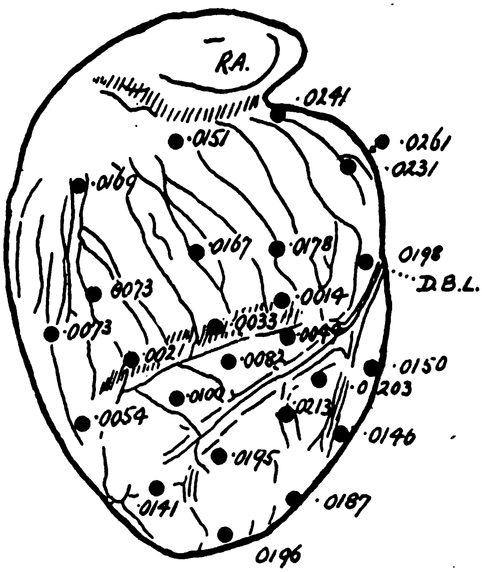


Fig. 56.—Diagram of experiment by Lewis showing the times at which the excitation wave appeared on the front of the heart relative to the upstroke of *R* in lead II. *R.A.*, right appendix; *D.B.L.*, descending branch of left coronary artery. (From Thomas Lewis.)

when nonpolarizable electrodes are placed at various parts of the outer aspect of the ventricle, and comparison made of the moments at which the cardiac impulse arrives, as judged by the appearance of the excitation wave relative to R in a standard electrocardiogram, it has been found that the time of arrival bears no relationship to the anatomic arrangement of the muscle bundles of the ventricle. It arrives early and simultaneously over an area of the surface near the anterior attachment of the wall of the right ventricle. It arrives late at the base of the right ventricle and in the part near the posterior intraventricular groove. Histologic examination has shown that the branches of the right division of the auriculoventricular bundle are most closely connected with the

place where the wave arrives earliest. Somewhat different results are obtained from the left ventricle, but again they are dependent upon the relationship of the part to the Purkinje fibers (Fig. 56).

FIBRILLATION OF THE HEART

Ventricles

The even spread of the wave of contraction over the heart depends on the uniform excitability of the muscular fibers. If certain of the muscular fibers, or bundles of fibers, have a greater or less excitability than others, then, when the stimulus to contract arrives, it will not produce a uniform contraction of neighboring bundles, and coordinated action of the cardiac musculature will give place to a confused movement in which each part of the heart is contracting independently of the rest. This *fibrillation,* or *delirium cordis,* as it is called, can be produced by a large variety of experimental methods, such, for example, as by stimulating the ventricles with induced electric shocks, or by ligation of a large branch of the coronary artery, or by the injection of lycopodium spores into the coronary circulation, or by mechanical stimulation of the heart in the region of the auriculoventricular bundle.

Fibrillation of the ventricles is undoubtedly a common cause of death in man, for of course the confused movements make the ventricles incapable of contracting on the contents of the heart. It is a condition which can probably never be recovered from in the higher animals, but it is of interest that the ease with which it is set up as the result of the application of an electric stimulus varies to a marked degree in different animals, and that in those hearts in which fibrillation can be elicited only with difficulty, recovery can usually be effected either by stopping the heart by means of cold and then allowing it to beat again, or by the administration of epinephrine. Of the hearts investigated in this way, that of the rat has been found to be most resistant to stimulation; then in order come those of the rabbit, the cat, the dog, and the horse. There is good reason to believe that the heart of man is readily affected. Fibrillation of the ventricle is undoubtedly the main cause of death in most cases of *electrocution.* Curiously enough, however, it has been stated that, whereas, a current of ordinary intensity (2300 volts alternating current) produces ventricular fibrillation in the heart of certain of the lower animals, at least in that of the horse, a very much stronger current does not do so, and may indeed cause ventricular fibrillation produced by a more moderate voltage to disappear. Unfortunately, however, these stronger currents produce irreparable damage in

the central nervous system, so that the method of applying stronger currents, even were it feasible to do so quickly enough, would be of no therapeutic value in removing fibrillation.

The disappointing results that have followed the repeated attempts to resuscitate persons killed accidentally by electric shocks is undoubtedly dependent upon the fact that in the heart of man it is impossible to bring back the normal beat after the ventricles have been thrown into fibrillation. Fibrillation of the ventricle is also the cause of the sudden cardiac failure occurring when blood clots or emboli cause a blockage of the coronary circulation (it is sometimes the cause of angina pectoris, for example). It must also be remembered in clinical practice that mechanical stimulation of the ventricles may produce fibrillation, so that in attempted resuscitation by cardiac massage care should be taken not to apply this too vigorously.

Auricles

Although ventricular fibrillation is seldom recovered from, it has been clearly shown in recent years that fibrillation of the auricles is relatively common and that it is by no means immediately fatal. Indeed it is one of the most common of the chronic cardiac disorders in man. Auricular fibrillation can be produced experimentally by the application of a strong electric stimulus to the auricles. If, however, a weaker stimulus is applied, the auricles do not go into typical fibrillation, but come to beat at a very rapid and regular rate, perhaps three or four hundred a minute. This condition is called "auricular flutter," and is quite frequently observed in the clinic.

The influence of auricular fibrillation and flutter on the beat of the ventricle is an extremely important one in connection with the irregularities of the heart observed in man, and this influence in most cases is explained by considering (1) the narrowness of the path (in the auriculoventricular bundle) along which the impulses have to travel, and (2) the varying conditions of excitability of the ventricular muscle, depending upon the existence of the refractory phase (page 180).

In auricular flutter, when three or four hundred impulses per minute are passing along the bundle to the ventricle, the contraction produced by the first one will scarcely have started before the second and immediately succeeding ones arrive, so that the ventricle will beat at a rate that is much less than that of the auricle, and a condition simulating heart-block will become established. The characteristic feature which distinguishes this from true heart-block, however, is the fact that the ventricular rate is *above* normal, whereas in true heart-block the rate is much below normal. By means of the electrocardiogram or by

polysphygmographic tracings, it can also be shown that the auricle is beating with perfect regularity although very rapidly.

In auricular fibrillation the ventricles obviously will respond at a very irregular rate to the impulses transmitted to them, and the auricular contractions, if examined by the methods above described, will show no regular sequence. Further details of the method of eliciting these signs will be described later (page 266).

CHAPTER XXIII

THE BLOODFLOW IN THE ARTERIES

THE PULSES

Returning to the physical laws that govern the circulation of the blood, we may now consider the temporary changes produced in the bloodflow in the arteries by each systolic discharge. These changes go under the general term of the pulses, of which three may be distinguished: (1) the pressure pulse, or the pulsatile increase of pressure produced by each heartbeat (see page 127); (2) the velocity pulse, or pulsatile acceleration of velocity; and (3) the palpable pulse, or the pulsatile expansion of the walls of the blood vessels produced by the sudden change of blood pressure in their interior. The general characteristics of the three pulses are the same, certain features being however more pronounced in one than in another.

General Characteristics

Rate of Transmission of Pulse Wave.—The rate of transmission of the pulse wave can be determined by taking simultaneous tracings of the pulses from two far distant parts of the arterial system along with accurate time-tracings. From records (cf. Fig. 98) taken from the apex or the carotid and radial arteries we can determine how long it takes for the beginning of the pulse wave to travel to the radial artery from the point in the aorta from which the carotid artery springs. We shall find that it takes about one-tenth of a second, which, considering the length of the artery involved, would work out as a transmission velocity of about seven meters per second or about seventeen miles an hour. The pulse therefore travels along the blood vessels at a much greater speed than the blood itself is moving, this being, as we shall see immediately, about 0.5 meters per second in the larger blood vessels.

The pulse is a wave of sudden increase in pressure and velocity passing along a stream which is flowing in the same direction with a certain more permanent pressure and velocity. A simple physical experiment may serve to make this clear: If the first of a row of billiard balls be tapped with the cue, the end balls will fly off while the others are moving slowly along in the direction of the stroke. Each ball becomes accelerated by the ball behind it, and imparts its influence to the ball

in front. In other words, a pulsatile acceleration of velocity is produced by a pulsatile change in pressure between each two balls. The existence of a pulse wave going in the same direction but quicker than a moving column of fluid can also be illustrated by observing the waves traveling down a stream when a stone is thrown into it.

The length of the pulse wave is such that the beginning of it has arrived at the periphery of the arterial system before the end has disappeared from the beginning of the aorta. This is important to remember, for it is a common mistake to think of the wave as being a local one. The determination of the length of the pulse wave depends upon the following equation: $L = VT$, where L equals the length of the pulse wave, V its velocity of transmission, and T its duration at a given point in the artery. Under ordinary circumstances L would usually work out from 3.25 to 4.5 meters.

The rate of transmission of the pulse wave varies according to the rigidity of the walls of the arteries. To understand why this should be so, it will be well for a moment to consider the physical conditions upon which the pulse wave depends. If we connect a piece of rigid tube with the nozzle of a large syringe, with each movement of the piston a wave of pressure will be transmitted to the fluid in the tube, along which it will travel at such a high velocity that it will arrive at the free end of the tube almost instantaneously, and incidentally the outflow of fluid from the end of the tube with each compression of the pump will be exactly equal to that represented by the movement of the piston. If, on the other hand, an elastic tube is employed, it will be found that the sudden increase of pressure produced by each stroke of the pump causes a distention of the walls, which travels along the tube as a wave at a readily measurable velocity, which is slower the more extensible the tube. Moreover, the outflow of fluid from the free end of the tube will continue for some time after the cessation of the movement of the pump. What happens in the tube with each discharge of the fluid is that the portion which is immediately adjacent to the pump undergoes distention and, being elastic, tends immediately afterward to recoil and thus exert a recoil pressure on the fluid contained in the tube. As a result, pressure waves are set up in the fluid in all directions. Those that travel back come to a stop because of the piston, while those that travel distally act on the fluid in front of them so as to accelerate it and by temporarily raising its pressure distend the next segment of the vessel wall, until the end of the tube is reached. From this consideration it is clear that the more extensible and elastic the wall of the tube is, the longer will it take for the wave of pressure to travel from one end to the other.

Alteration in the rate of transmission of the pulse wave in the arteries of man depends entirely upon an application of these principles. When the arteries become hardened in old age, the rate of transmission of the pulse wave is markedly increased. The pulse is also transmitted more rapidly in the vessels of the lower extremities than in those of the upper, since in the former the blood vessels are somewhat more rigid. Delay in the transmission of the pulse wave is further observed as one of the signs of aneurism in a vessel; as is well known, aneurism of the subclavian artery on one side causes a delay of the pulse on that side that is perceptible to the fingers.

The Contour of the Pulse Curves

For more particular study of the exact contour of the pulse wave, and especially for determining the time relationships of the secondary waves,

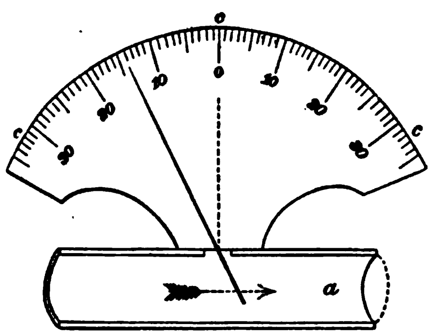


Fig. 57.—Diagram of Chauveau's dromograph. *a*, tube for introduction into the lumen of the artery, and containing a needle or vane, which passes through the elastic membrane in its side and moves by the impulse of the blood current; *c*, graduated scale for measuring the extent of the oscillations of the needle.

a large variety of methods of varying degrees of accuracy have been elaborated for each kind of pulse.

Those devised for measuring the *pressure pulse* have already been described (see page 127), and for the other pulses they are as follows:

Velocity Pulse.—Much ingenuity has been displayed in the elaboration of methods for recording the velocity pulse. In one of these the artery is cut across and the ends attached to a tube, into the lumen of which projects a paddle or vane articulated with a light lever, which passes through its wall (see Fig. 57). The vane floats in the blood stream, and the outer end of the lever to which it is attached is connected with some device to record its movements, which vary with the velocity of bloodflow (hemodromograph). Another method consists in the application of the instrument known as Pitot's tube used by physicists. This consists of a horizontal tube having two side tubes, each of

which is connected at its outer end with a manometer and prolonged inside the horizontal tube, where they are bent at opposite right angles, so that the inner end of one of them—the proximal tube—points up

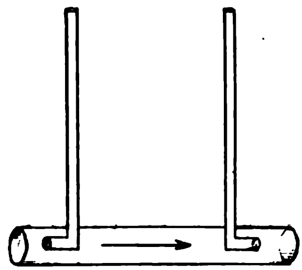


Fig. 58.

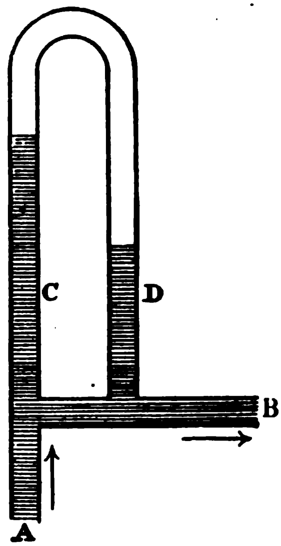


Fig. 59.

Fig. 58.—Diagram to show principle of Pitot's tubes for measuring velocity pulse. In both tubes the fluid will rise because of lateral pressure, but in the proximal (left-hand) tube it will rise higher than in the distal, because it will also be affected by the velocity of flow.

Fig. 59.—Diagram to illustrate the principle of Cybulski's Photo-hematotachometer. The fluid in *C* stands higher than that in *D* in proportion to the velocity of flow of the blood along *AB*.

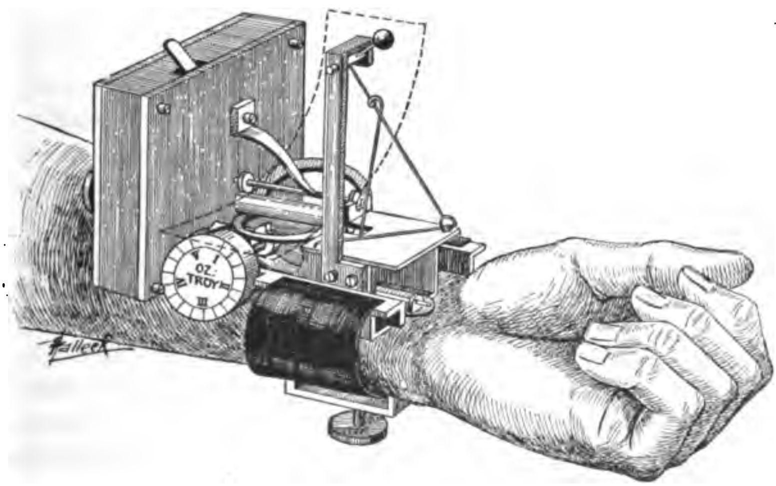


Fig. 60.—Dudgeon's sphygmograph. (From Jackson.)

stream, and records not only the lateral pressure but also the pressure produced by the sudden increase in velocity of the flow, while the

other—the distal tube—being bent down stream, records merely lateral pressure. A photographic record of the movement of the fluid in the two tubes gives the velocity pulse (see Fig. 58). For physiologic purposes the form of apparatus used is constructed as shown in Fig. 59.

Palpable Pulse.—To secure a record of the palpable pulse, the so-called sphygmograph is employed, although a tambour having a button in the center which is made to press on the artery may also be employed. The commonest form of sphygmograph is that known as Dudgeon's (Fig. 60). It consists of a small button connected with a spring, the movements of which are transmitted and magnified by means of a system of levers connected with a writing point arranged so as to inscribe its movements on a moving surface.

The Analysis of the Curve

The general contour of the pulse waves taken by any of the above methods are in general very much the same. The pressure and velocity

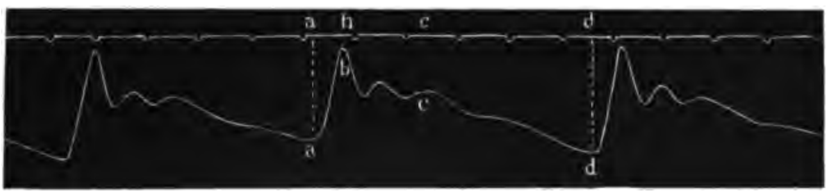


Fig. 61.—Pulse tracing (sphygmogram) taken by sphygmograph. *a d*, the period of the pulse curve; *b*, the primary; *c*, the dicrotic wave. Time marked in fifths of a second. (From *Practical Physiology*.)

pulse curves are, however, not usually taken for the purpose of observing the contour of the wave but rather for measuring the difference in pressure or velocity actually produced during each pulse. It is a record of the palpable pulse that is usually employed for studying the contour of the wave and the presence of secondary waves. A record of the palpable pulse wave (Fig. 61) shows two separate waves on the descending limb of the main wave. If a large number of similar pulse curves are taken from different individuals or from the same individual under different conditions, it will be found that of these two waves the second one is by far the more constant; and if the waves are timed in relationship to the heart sounds, this second wave always occurs immediately after the second sound, allowance, of course, being made for the time required for the pulse to be transmitted from the heart to the artery from which the pulse tracing is being taken. If the observation is made very carefully, it can indeed be shown that the second sound corresponds exactly to the notch which precedes this wave. The waves that

precede this notch can not be related to definite changes occurring in the heart. Evidently, then, the secondary pulse waves are not all of equal significance, by far the most important being that which occurs immediately after the second sound, called the *dicrotic wave* (c), the notch in front of it being called the *dicrotic notch*. Any secondary waves occurring before the dicrotic are called *predicrotic,* or if they occur on the ascending limb of the main pulse wave, as they sometimes do, they are called *anacrotic*. Waves occurring after the dicrotic are called *postdicrotic*.

The relative importance of the dicrotic, in comparison with the predicrotic and postdicrotic waves, is further evidenced by the fact that it alone is seen on a so-called *hemataugram,* which is the tracing obtained by allowing a fine stream of blood, escaping from a pinhole made in the wall of an artery, to impinge upon a moving sheet of white blotting paper. That such a tracing shows a dicrotic but no secondary wave, indicates that only the former is present in the blood stream itself, and that the other secondary waves must be produced by some condition arising either in the elastic tissue of the walls of the blood vessels, or in the elastic properties of the instruments used for taking the pulse tracing.

The Dicrotic Wave.—Because of its obviously greater significance, we shall first of all consider the exact cause of the dicrotic wave and of the notch preceding it. Theoretically, two possible causes might explain the wave: either it is due to some secondary wave set up at the heart, or it is dependent upon waves reflected from the periphery of the circulation back along the blood stream, just as secondary waves are reflected from the walls of a tub of water when a stone is thrown in the center. In considering this second possibility, we are of course making the assumption that at the ends of the arterial system there is a sudden resistance to the onward movement of blood. The frequent branching which occurs when the arterioles open into the capillaries no doubt offers many opportunities for the reflection of pulse waves back to the heart, but these waves must be reflected at such varying distances along the arterial system that there can be little opportunity for them to become added together so as to form a wave of sufficient magnitude to make itself perceptible in the blood flowing in the larger arteries. These waves are relatively so small and they occur at such different times that the net result of their addition, so far as the production of a larger wave is concerned, must be practically nil. Notwithstanding these considerations, it is possible that under some conditions, such as in cases of high arterial tension, certain of the predicrotic or postdicrotic waves may be due to the above causes.

That *the dicrotic is not a reflected wave* is clearly established by the fact that if the distance between the dicrotic wave and the main pulse wave is measured at different points of the arterial stream, it will always be found to be the same, which obviously would not be the case were the dicrotic wave reflected. If, for example, we were to examine the contour of the wave produced by throwing a stone into a tub of water, we should find that near the edge the secondary wave was very close to the main wave, whereas near the center the secondary wave would occur much later.

Our problem therefore narrows itself down to an investigation of the cause for the dicrotic wave at the central end of the circulation. It occurs, as we have seen, immediately after the beginning of diastole. That it can not be due to anything taking place in the ventricle itself is evidenced by the fact that such a wave is absent from an intracardiac pressure curve (see page 151), although it is present in the very beginning of the aorta. Now, the only structures existing between those two points which could be held responsible for this wave are the semilunar valves—a conclusion which is sustained by the fact that, if the aortic valves are rendered incompetent by hooking them back, or if the pulse beat is examined in patients suffering from an aortic insufficiency, it will be found that the dicrotic wave is not nearly so evident as usual.

To understand *how the valves are responsible for the production of the wave,* the mechanical changes occurring at the root of the aorta must be clearly understood (see page 155). The stretching of the elastic walls of the aorta which occurs with each systolic outrush of blood is followed by a powerful and sudden contraction of the stretched walls, and the pressure thus brought to bear on the column of blood in the aorta tends to impel it both forward and backward. The forward movement adds itself to the wave of increased pressure already produced by the ventricular contraction. The backward component travels as far as the semilunar valve, from which it is reflected, and now proceeds peripherally along the blood stream during the time at which the original pressure pulse is declining. It therefore imprints itself on the pulse tracing as a separate wave, and does so all the more markedly when the decline in the main pulse wave is rapid, as in cases in which the peripheral resistance is low, but fails to be prominent when, on account of a high peripheral resistance, the decline in the main pulse wave is tardy. This explanation coincides exactly with the well-known clinical fact that the dicrotic wave is conspicuous in pulses of low tension, but ill marked or absent in pulses of high tension.

One point remains to be considered, and that is *the cause for the sudden decline in the main wave* at the cessation of the ventricular out-

put, for, it might be said, why should there be such a sudden fall in pressure near the heart, whereas toward the periphery, as we have seen, the pressure between the heartbeats tends to be maintained on account of the elastic recoil of the stretched arterial walls. The explanation usually given is that the sudden cessation of outflow of blood from the ventricle at the end of the sphygmic period causes a negative pressure to be produced in the blood at the beginning of the aorta, thus tending to cause a reflux of blood towards the heart, the effect of which is (1) to bulge the closed valves, and (2) to produce the reflected dicrotic wave. If, while fluid is flowing under pressure along a tube, the flow is suddenly arrested by turning a stopcock, it is possible by the use of manometers to show that a negative wave is set up immediately beyond the stopcock, and that this negative wave travels along the tube at a rate depending on the elasticity of its walls.

Causes for Disappearance of the Pulse in the Veins

The disappearance of the pulse in the capillaries and its consequent absence in the veins we have already seen to be owing to the combined influence of the elasticity of the vessel walls and the peripheral resistance. On account of these two factors the pressure conveyed to the blood during systole is stored up to be released during diastole by the recoil of the stretched vessels. Sometimes, however, the pulse gets through to the veins, either because the elasticity of the vessels is not so marked, or because the peripheral resistance has been lowered (vasodilatation). In patients with hardened arteries, or in normal individuals after taking nitrite, which dilates the peripheral arterioles, a pulse may come through at the periphery and appear in the veins. This may be called the peripheral venous pulse, and it is to be carefully distinguished from the central venous pulse observed in the large veins, as at the root of the neck, before any valves have intervened to block the transmission of the auricular pressure wave back into the column of blood in the veins. If a pulse is seen in a large vein and there is doubt as to whether it is peripheral or central in origin, this doubt can be immediately removed by locally constricting the vein; if the pulse is peripheral, it will disappear on the heart side of the constriction; if it is central, on the side away from the heart.

CHAPTER XXIV

THE RATE OF MOVEMENT OF THE BLOOD IN THE BLOOD VESSELS

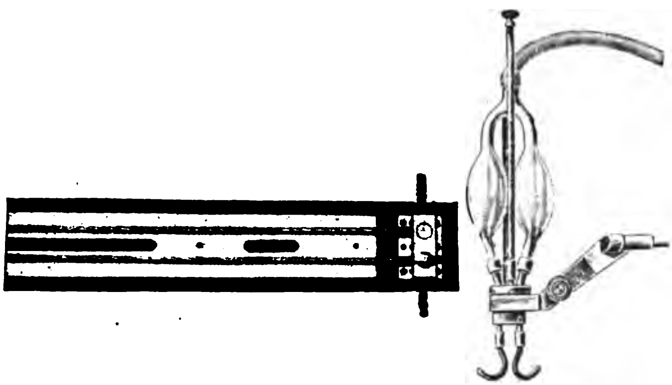
Since the object of the circulation is to maintain an adequate movement of blood in the tissues and capillaries, it is evident that besides measuring the pressure of bloodflow, we should also measure the rate of its movement, or, as it is often called, the mean velocity. This measurement may be undertaken either for a given vessel or for a complete vascular area, such, for example, as that of one of the viscera or one of the extremities—the mass movement of the blood. Or instead of measuring the mean velocity we may desire to know how long it takes for a particle of blood to traverse a given vascular area. Such a measurement is called the *circulation time;* it does not at all tell us how long it takes for *all* the blood to pass through the given area, but only, as stated, the time required for the circulation of a fraction of the blood through a particular field.

VELOCITY OF FLOW IN A VESSEL

Special methods have been devised for the measurement of each of these three velocities. For the measurement of the velocity of flow through a main artery or vein, methods similar to those employed by hydraulic engineers are employed; that is to say, the volume of blood, in cubic centimeters, which passes a given point is measured for a given time, and the result divided by the cross section of the vessel at the point of observation. The result gives us the *mean lineal velocity.* To measure the outflow of blood in a given time, the simplest method would be to cut across the vessel and collect the blood in a graduate, but obviously in this method an error would be introduced, because cutting the vessel would lower the peripheral resistance and remove the natural obstruction to the flow present in the intact animal. Moreover, the hemorrhage would in itself introduce a disturbing factor on account of the loss of circulating fluid.

To make such measurements of any value, it is obviously necessary to retain the peripheral resistance. For smaller vessels this can be done by introducing in the course of the artery a long glass tube bent in the

shape of the letter U (Fig. 62), or by merely allowing the vessel to bleed into a graduated tube and seeing how long the blood column takes to travel from one end to the other. This method is of considerable value in measuring the velocity of flow from small vessels such as the veins coming from glands and muscles. For larger vessels a so-called *stromuhr* is employed. There are numerous forms of stromuhr; that shown in the diagram (Ludwig's) (Fig. 62) consists of two glass bulbs united above, and connected below with tubes that open flush with the surface of a brass disc. This is pivoted at its center with another similar platform also having flush with the surface the openings of two tubes connected with the cut ends of the artery or vein. In a certain position of the platform, the tubes from the artery or vein are exactly opposite those of the bulbs, so that the blood can flow from one end of the vessel



1. *2.*

Fig. 62.—Forms of apparatus for measurement of blood velocities.
1. Volkmann's hemodromometer. The blood vessel is attached to the two short side tubes, and according to the position of the stopcock, the blood flows either directly between them or through the U-shaped glass tube.
2. Ludwig's stromuhr. The tubes on the lower end of each of the two glass bulbs pierce a circular brass platform and end flush with its surface. This platform pivots at its center on a similar lower platform with two openings connected with the tubes that lead to the blood vessel.

through the bulbs to the other end. To use the instrument the proximal bulb is filled with oil and the peripheral one with physiologic saline. The clip is then removed from the central end of the artery, and the blood flows in and displaces the oil, which in turn displaces the saline in the peripheral end of the artery. When the blood has risen to a mark on the tube joining the two bulbs, the instrument is rapidly rotated so that the oil is brought back again into the proximal position, the rotation being effected so quickly that there is no distinct interruption in blood-flow. The operation is repeated in this way for a given period of time. Counting accurately the number of revolutions, then multiplying the number of revolutions by the capacity of the bulbs, we get in cubic

centimeters the amount of blood that has flowed through the instrument in a definite unit of time. This gives us the volume flow and, if the result is divided by the cross section of the vessel in square centimeters, we obtain what is known as the mean lineal velocity. Many modifications have been made of this instrument, but it is unnecessary to go into them here.

The general result of such measurements has been to show that the lineal velocity is *inversely proportional to the cross section* of the vessel at the point of observation. It is obvious that the volume of blood flowing out of the heart to the aorta in a given time is exactly equal to that flowing into it by the vena cava, and likewise that the volume flowing into an organ is exactly equal to that which flows out. Consequently the lineal velocity will be inversely proportional to the sectional area of the vessel. The principle is the same as that which governs the velocity of flow of a stream: when the bed is narrow, the current is swift, but it becomes sluggish when the bed is wide. If the arteries were of the same caliber as the veins, the mean velocity of the bloodflow through the two would be the same, but actually it is much greater in the arteries because the lumen of these at a given point in the circulation is only from one-third to one-half that of the corresponding vein.

It must be understood that we are dealing above with the mean velocity in a unit of time, and that there must be considerable alteration with each systole and diastole, constituting the velocity pulse (page 200). The degree of this alteration with each velocity pulse is much less at the periphery of the circulation than near the heart. As the periphery is reached, the flow becomes more uniform. It must further be remembered that, although the mean velocity depends essentially upon the area of the vascular bed, yet it is subject to considerable variations as a result of changes either in the force or rate of the heartbeat or in the facility of outflow from the ends of the arterial system—that is, changes in peripheral resistance.

It is usually stated that the mean lineal velocity in the carotid artery is about 300 millimeters per second; and in the jugular vein, about 150 millimeters; whereas in the capillaries, where the total area of the vascular bed has become enormously increased, being perhaps some 800 times that of the aorta, the velocity of flow is only about half a millimeter per second.

MASS MOVEMENT OF THE BLOOD IN A VASCULAR AREA

Methods.—In considering *the bloodflow or mass movement of the blood* in the different regions of the body, it is usually more practical to

measure, not the mean lineal velocity of the inflowing and outflowing blood, but rather how many cubic centimeters of blood are traversing the part per 100 grams of organ or tissue per unit of time. Such measurements may be made in a variety of ways. If there are but one artery and one vein to the part, the stromuhr may of course be employed, and it may be inserted in either the arterial or the venous circuit. For measuring the mass movement of blood through such large viscera as the liver, this is indeed the only method that can be employed, the stromuhr being inserted either in the course of the portal vein and hepatic arteries, or, better still, in the vena cava just below the openings of the hepatic vein, the vena cava being shut off for a moment between the liver and the heart and the blood, as it flows from the hepatic vein, allowed to collect in the stromuhr. For other organs and tissues, however, methods which do not involve any interference with the blood vessels may be employed. One of these is the so-called *plethysmographic method of Brodie.* An organ, such as the kidney, is enclosed in a plethysmograph (see page 230,) and while a record of its volume is being inscribed on a quickly revolving drum, the vein is suddenly clamped, with the result that the kidney volume expands in proportion to the mass of blood flowing into it. When the expansion has reached a certain degree, the clamp is removed and the bloodflow allowed to pursue its course. It is then an easy matter, by graduating the plethysmograph, to determine how many cubic centimeters of blood must have flowed into the organ in the given time. To avoid serious local asphyxia in the tissue, the clamp must be applied to the vein for only the briefest period of time. This method may also be employed for measuring the bloodflow through the extremities. Thus, if the arm is enclosed in the plethysmograph (Fig. 63) and a band encircling the arm above the plethysmograph is tightened so as to constrict the veins but not the arteries, the rate at which the volume of the arm within the plethysmograph expands will correspond to the rate at which blood is flowing into it (Hewlett).

For the purpose of measuring blood flow through the upper or lower extremities, a much more serviceable *clinical method* is that of G. N. Stewart. This depends on the principle that, provided the blood passing from the thorax to the hands or feet is of constant temperature, the rate at which heat is dissipated from the hands or feet will be directly proportional to the rate of movement of the blood through these parts. Fortunately for the method, the hands particularly, but also the feet, are more or less perfect radiators—at least they are to this extent, that if the temperature in their environment is not much lower than the temperature of the blood, then while this is traversing the part, it will

lose heat to the environment until the outflowing or venous blood is at exactly the same temperature as the environment; for example, if the hand is placed in water that is a little cooler than that of the blood, and the temperature of the blood in one of the large veins of the hand is measured, it will be found to be the same as that of the water in the water-bath.

To measure the rate of flow, therefore, we must ascertain: (1) how much heat has been given out by the part to the water surrounding it in a given time, and (2) the difference in temperature of the inflowing (arterial) and outflowing (venous) blood. We measure the amount of

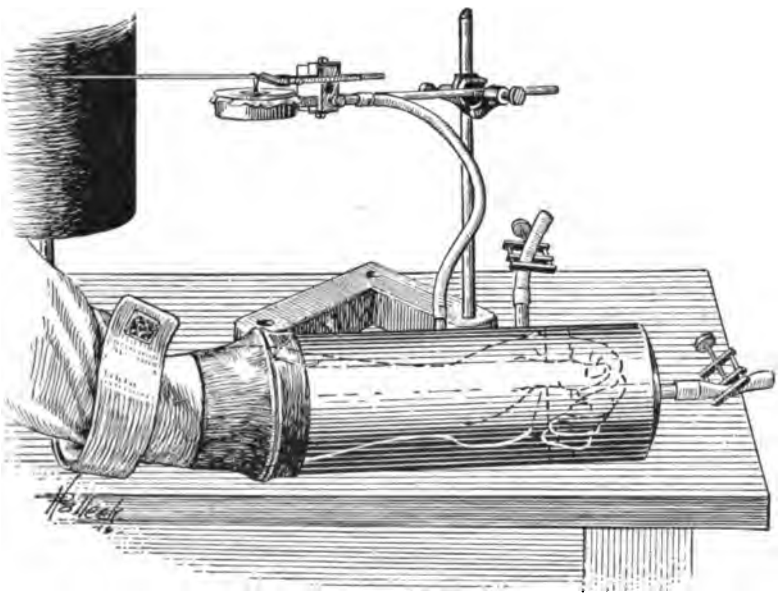


Fig. 63.—Plethysmograph for recording volume changes in the hand and forearm. By observing the rate with which the volume increases when the arm is compressed, the mass movement of the blood can be determined. (From Jackson.)

heat given out to the water in calories, a calorie being the amount of heat required to raise the temperature of 1 c.c. of water from 0° C. to 1° C. Suppose, for example, a hand were placed in 3,000 c.c. of water at 33° C., and that after ten minutes the temperature had risen to 33.5° C., then the amount of calories given out would be 3,000 x 0.5= 1500. Since calories equal cubic centimeters multiplied by change in temperature, it follows that if we divide the figure representing them by the *actually* observed difference in temperature between inflowing and outflowing blood, the result must equal the number of cubic centimeters of blood that has flowed through the part. The temperature of the inflowing blood has been found to be practically identical with that of the

mouth under the tongue; whereas of course the temperature of the venous blood, as already explained, is equal to the mean temperature of the water during the time that the hand was immersed in it. Further details of the technic of this method will be found elsewhere, but it may be said here that it is extremely simple and accurate, and that it requires nothing more than (1) an accurate thermometer ranging between about 40° C. and 50° C., with a scale so drawn out that readings can be made to $\frac{1}{100}$ of a degree, and (2) a well-constructed vessel of about 3,000 c.c. capacity, with double walls, the space between them being packed with some heat-insulating material such as ground cork.

Results.—Regarding the results obtained with these methods, it has been found that the blood supply for each 100 grams of tissue in the viscera, as measured by the stromuhr method, is about as follows: stomach, 21 c.c.; intestine, 71 c.c.; spleen, 58 c.c.; liver, arterial, 25 c.c.; liver, venous, 59 c.c.; liver, total, 84 c.c.; brain, 136 c.c.; kidney, 150 c.c.; thyroid gland, 560 c.c. The large blood supplies of the thyroid gland and of the kidney are the most striking results of these observations.

By the use of the calorimeter method the bloodflow through the hands and feet of a healthy young man has been found to be about 13 grams per 100 c.c. of hand per minute for the right hand, and about half a gram less for the left. The footflow is only about one-third to one-half that of the hand per 100 c.c. of tissue—a difference which is largely owing to the greater proportion of skin and the smaller proportion of bone in the hand. The average footflow or handflow for a given individual under ordinary conditions is remarkably constant from time to time, but it is extraordinarily sensitive to changes in the temperature of the environment in which the subject has been living for some time previous to the measurement. In one individual, when the room temperature was 20° C., the flow in the right hand, expressed in grams of blood per 100 c.c. of hand or foot, was 10.1; when it was 22.8° C., the flow was 12.8; when it was 25° C., 12.1; when it was 30° C., 18.5. On account of the influence of temperature on the flow, it is extremely important that the measurements should be made in a small room the temperature of which is kept constant, or if it must be made in the wards, the bed should be surrounded by curtains. The measurements made on the hands of dispensary patients shortly after coming in from outside air are very likely to be fallacious. The importance of making such bloodflow measurements in the clinic will be alluded to later.

Of course the measurements made by the above method in man tell us only the rate of flow in the periphery of the body, and furnish us with no information regarding the flow of blood through the viscera. It is, however, a well-established fact that the bloodflow in the central part of the

circulation is more or less reciprocal with that at the periphery, an increase in the one place being accompanied by a corresponding decrease in the other.

The Visceral Bloodflow In Man

The visceral bloodflow in man can be measured indirectly in the case of the lungs, either, (1) by finding the quantity of oxygen absorbed by the blood during an interval of time that is less than that required for the blood to travel once round the circulation (60 seconds) and comparing this with the oxygen content of samples of arterial and venous blood, or (2) by causing a person to breathe a known quantity of nitrous-oxide gas and then finding the concentration of this gas in the blood after leaving the lungs. In the former method the difference in oxygen percentage between arterial and venous blood will be less for a given absorption of oxygen from the alveoli the more rapid the circulation of blood through the lungs, and in the latter method for the absorption of a given amount of nitrous oxide, the less will be the concentration of this gas in the blood the more rapid the circulation. Obviously these estimations must be made only over periods of time, less than that taken for any of the blood to complete one circuit of the circulation.

The methods are admittedly only approximate, but the results are of much interest, mainly because of the indication they give us as to the amount of blood pumped out by the ventricle with each heartbeat, or during a given period of time. The results have been found to vary considerably; thus, one author (Krogh) places the output of blood per minute as between 2.8 and 8.7 liters, which would correspond, at a pulse rate of 70, to an output per heartbreat of from 40 to 120 c.c. An immediate and very marked increase has been found to occur during muscular work. By comparing the bloodflow through the hand with that through the lungs, an estimate can be formed in a given individual as to the relative magnitude of the peripheral and visceral moieties of blood. Interesting results, which will be referred to later, have been obtained from such measurements.

The Work of the Heart

Meanwhile it is of interest to note that we may calculate from the ventricular output of the blood *the amount of work that the heart is doing in maintaining the circulation.* Of course the calculation is again only approximate, since we have to assume certain figures. If we assume that in a 70-kilogram man the quantity of blood is 4,200 c.c. (see page 85), and that it takes about one minute for all the blood to complete a circulation, then the work performed by the left ventricle in one minute

will be equal to that done in raising the above quantity of blood to a height corresponding to the mean pressure in the aorta. If we take this pressure as 130 millimeters of mercury, which would correspond to a column of blood 1,755 meters high (13.5 x 130=1755 mm. blood, or 1.755 meter), the work done by the left ventricle would be 1.755 x 4.2=7.37 kilogram-meters in one minute, or in twenty-four hours roughly about 10600 kilogram-meters. The work done by the right ventricle is probably about one-third that of the left, this being about the ratio of the pressures in the two chambers. The total work of the two ventricles is therefore about 14000 kilogram-meters. This represents an enormous amount of work; indeed it has been computed that it is sufficient to raise a man of 70 kilograms to about twice the height of the highest skyscraper in New York. The work thus expended in forcing the blood through the capillaries becomes converted by friction in the small blood vessels into heat, the heat equivalent of the above amount of work being roughly about 350 calories (see page 537).

THE CIRCULATION TIME

The circulation time, or the time taken by a drop of blood to travel between two points in the circulation, can be determined in laboratory animals by a variety of methods, all depending on the principle of seeing how long it takes for a drop of some substance injected into an artery to appear in the corresponding vein. For example, to determine the time taken for a drop of blood to pass from the jugular vein into the carotid artery in a rabbit, a solution of methylene blue in isotonic saline is injected into the former vessel and the moment of its appearance through the walls of the artery determined by a stop-watch. If the walls are too thick to admit of the employment of this method, a strong solution of sodium chloride may be substituted for the methylene blue, and the moment of its appearance at another point of the circulation determined by observing the electrical conductivity of the vessel. Since the conductivity of a blood vessel depends partly on the concentration of electrolytes in the blood flowing through it, the moment at which the salt solution appears will be indicated by a change in electrical resistance (G. N. Stewart).

By such methods, it has been found that the time for the pulmonary circulation is very short compared with that of the systemic circulation. In a rabbit it is usually a little less than four seconds; in an average-sized dog of about 12 kilograms, it is about eight seconds; and in man it is computed to be about fifteen seconds. On the other hand, the circulation time in such viscera as the spleen and kidney is relatively long,

and more susceptible than that of the lungs to different conditions of temperature. In a dog in which the pulmonary circulation time was about 8.5 seconds, that of the spleen was about 11 seconds, and of the kidney about 17.5 seconds. The shortest circulation time of all is of course that in the coronary artery, but that through the retina can not fall far behind it.

To determine *the total circulation time,* we must know: (1) the average amount of blood passing by each part in a given time, and (2) the average circulation time of each part. From such computations, which however are obviously subject to considerable error, it has been reckoned that the total circulation time in man must lie somewhere between 1 and 1.25 minutes.

MOVEMENT OF BLOOD IN VEINS

Before leaving this part of our subject, a few words may be said concerning *the forces concerned in the movement of blood in the veins* from the capillaries to the heart. By the time that the venules are reached, owing to friction in the capillaries the blood will have lost most of the force imparted to it by the heart action. Nevertheless, this remaining *vis a tergo* must be considered as the basic cause for the movement of the venous blood near the periphery. As the venules get larger, two other factors come into play: the massaging influence of the muscles, and the valves of the veins. By the movements of the muscles the veins which lie between will be rhythmically compressed, and this will tend to cause the blood to be moved forward and backward in them, the backward movement being however prevented by the operation of the valves. When the tonicity of the muscles is subnormal, as in conditions of ill health, the absence of this massaging action permits the blood to stagnate in the veins, especially in those of the lower extremities in upright animals, with the consequence that the veins become dilated, particularly just above the valves, thus causing the condition known as varicose veins.

As the thorax is approached, two other factors become operative: the aspirating influence of the thorax during inspiration, and the negative intraventricular pressure (see page 152). There is no doubt that the former of these is of considerable importance in maintaining the venous return near the heart, for although the change of pressure induced by inspiration amounts to only 5 millimeters of mercury, yet it acts so slowly that it produces a considerable influence. The aspirating effect of the ventricle at the beginning of diastole is, however, of no significance in attracting blood to the heart, for although, as we have seen, it may be considerable, yet it lasts for so short a time that it could not

overcome the inertia of the column of blood in the vena cava. Even if the negative pressure did last for a longer period, it could not attract more than a small amount of blood, because it would cause the thin collapsible walls of the veins to come together and thus block the passage towards the heart.

CHAPTER XXV

THE CONTROL OF THE CIRCULATION

The available blood in the body is parceled out to the various organs and tissues according to their relative activities, and, since these vary from time to time, the question arises as to the nature of the mechanism or mechanisms involved in bringing about this adjustment. Two possible methods of increasing the supply are: *an increase in the mass movement of all the blood* in circulation, and *a reciprocal adjustment of the resistance to the flow in different vascular areas* brought about by vasodilatation in one and vasoconstriction in others. Both of these methods might operate together.

Two agencies can be thought of as responsible for bringing about the above changes: (1) chemical substances or hormones, present in the blood, and (2) the nervous system.

The influence of chemical substances, or hormones, (page 729) *in the control of the circulation* is undoubtedly an important one, and of those known at the present time two groups may be mentioned: (1) substances which alter the hydrogen-ion concentration of the blood, and (2) so-called pressor and depressor substances, produced either by ductless glands, such as the adrenal, or by the activity of tissues. An increase in hydrogen-ion concentration of the blood not only affects the heartbeat (see page 168), but causes a marked dilatation of the blood vessels, so that both the central and the peripheral changes will be such as to encourage an increased flow of blood through the active organs or viscus. Thus, during muscular activity of the leg muscles there will be a tendency to an increase in the hydrogen-ion concentration of the blood as a whole, resulting in a greater cardiac activity and a greater outrush of blood through the aorta, and at the same time the vessels of the acting muscle will have become especially dilated because of the production by the active muscles either of lactic acid or of carbonic acid. The active muscle also produces such substances as imidazole, which have a powerful vasodilating action. Such substances are sometimes called depressor.

Though the hormone control of the circulation is undoubtedly of great importance, it is probably much less so than that exercised through *the nervous system,* and here again the control is centered partly in the

heart and partly in the peripheral resistance. The nerve control of the heart is effected through the vagus and sympathetic nerves, and that exercised on the blood vessels, through the so-called vasoconstrictor and vasodilator nerves.

The activity of the nerve centers from which the cardiac and vasomotor impulses are discharged is controlled by afferent impulses coming from the various regions of the body. When a gland becomes more active, we must suppose that stimulation of the sensory fibers has caused afferent impulses to be transmitted to the cardiac and vasomotor centers, upon which they act in such a way as to produce increased heart action and a local dilatation of the blood vessels of the active gland, with perhaps a constriction of the blood vessels of the rest of the body.

THE NERVE CONTROL OF THE HEARTBEAT

The Vagus Control

With regard to the control exercised through the vagus nerve, we have already seen that the cutting of the two nerves in the neck causes the heart to quicken and the arterial blood pressure to rise, whereas a stimulation of the peripheral end of the nerve causes the heart to become slowed, if not stopped altogether, and the blood pressure to fall.

For the more detailed investigation of the nature of the vagus control of the heart, it is necessary to observe the exposed heart itself—an experiment which, for obvious reasons, can be most simply performed in a cold-blooded animal, such as the frog or turtle, but which can also be performed in mammals provided artificial respiration is maintained. The general effect of the vagus in both groups of animals is the same, although apparent differences may exist on account of the relative importance of the different parts of the heart in the origination and propagation of the heartbeat.

The Cold-Blooded Heart.—If the vagus nerve on the right side in the turtle (the left nerve is inactive in this animal) is stimulated with a very feeble electric current, while simultaneous records are being taken of the contractions of the auricles and ventricles in the manner shown in the accompanying tracing (Fig. 64), it will often be found that there is a weakening of the auricular beats without any change in those of the ventricle. If the strength of stimulus is somewhat increased, the auricular beat, besides becoming weaker, will also become slower, but meanwhile the ventricular, although also slower, may become distinctly stronger. At first sight this result may be a little confusing, because it would seem to indicate that the vagus nerve weakens the auricular,

but strengthens the ventricular beat. It is clear, however, that the strengthening of the ventricular beat is merely due to the fact that the cavity has become better filled with blood during diastole as a result of the slowing of the auricle. These results indicate, then, that with weak stimulation the vagus exerts its direct influence only on the auricle. If

Fig. 64.—Simultaneous tracings from auricle and ventricle of turtle's heart. Between the crosses the vagus was stimulated, with the effect that the auricular beat diminished in force but not in frequency, while the ventricular beats were practically unaffected. (From Howell's *Physiology*.)

the stimulation is strong enough both auricles and ventricles cease to beat altogether, and if the stimulus is maintained, the inhibition may go on for a very long time (Fig. 65).

Usually, even though the stimulus is maintained the heart begins to

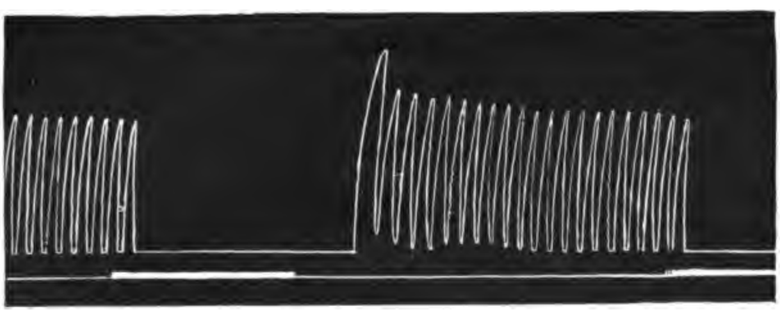


Fig. 65.—Effect of vagus stimulation on heart of turtle. Note the after effect of augmentation.

beat again after a time, at first only occasionally but gradually more rapidly. This is known as *escapement,* and it indicates that the energy pent up in the heart during the vagus inhibition has at last overcome the inhibiting influence of the nerve, which is meanwhile becoming fatigued. All of these results could be quite satisfactorily explained on the assumption that the action of the vagus is confined to the sinus,

which, it will be remembered, dominates the beat in the rest of the heart. There is evidence, however, that the vagus also directly affects the rhythm of the ventricle. It may be stated as a general conclusion from these results that *the influence of the vagus upon the heartbeat is chiefly centered upon those parts of the organ in which the rhythmic power is most highly developed.*

Besides affecting the rate and strength of the heartbeat, the vagus also exercises a control on the conductivity of the cardiac muscle. Thus, if a partial block is instituted in the turtle heart by applying a clamp between the auricles and ventricles, stimulation of the vagus enfeebles the auricular beat and may also cause a complete heart-block as shown in the tracing reproduced in Fig. 66. It is important to point out here, however, that under certain conditions the vagus may appear to increase rather than decrease the conductivity of the tissue in the auriculoven-

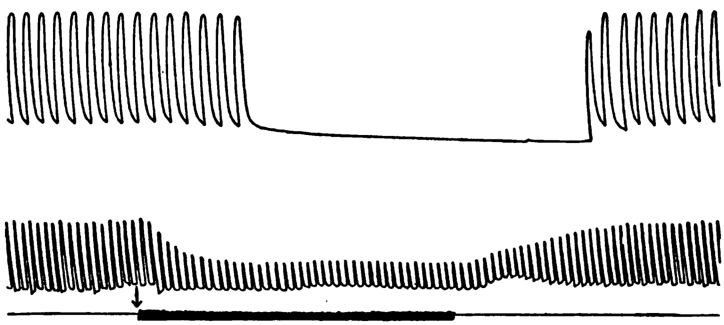


Fig. 66.—Tracing to show that vagus stimulation may diminish transmission from auricles to ventricles. It shows the effect of stimulating the left vagus on partial (2/1) block produced on heart of turtle by application of clamp at auriculoventricular junction. Stimulation at ↓ depressed the conductivity and weakened the auricular contractions (lower tracing) without slowing their rate. The result was an increase in the degree of block with cessation of ventricular contractions (upper tracing). Initial auricular rate = 35 per minute. (From Garrey.)

tricular junction; for example, it has been observed in the turtle heart that when a clamp is so tight as to produce complete block—that is to say, to render the ventricle inactive while the auricle is still beating at the usual rate—stimulation of the vagus, besides causing the auricles to become distinctly slowed, may at the same time cause the ventricles to respond to the auricular beats. This result is probably due to the better chances of slow beats getting through the junction than those which are so frequent as to crowd one another on the narrow bridge which the constricted tissue offers to their passage (Fig. 67).

Very important work was contributed in this field by G. R. Mines[13] shortly before his lamentable death. He found that the local application of atropine to the sinus eliminates the effect of stimulation of the (intracranial) vagus on the rate of the heartbeat, while the effect on the

auriculoventricular junction and on the ventricle remains. After the atropinization, vagus stimulation delays the transmission of beat from auricle to ventricle and shortens the time of each beat in the ventricle. It was further found that by the local application of atropine various parts of the ventricle can be rendered irresponsive to the influence of the vagus and the effects of this nerve on the form of the cardiogram modified at will. These results have an important bearing in the interpretation of the cause of the T-wave of the electro-cardiogram which will be referred to later. Mines' results show that the probable explanation is that the T-wave is due to the greater duration of the excitatory state at the base than at the apex, for by altering the relative duration of this state at base and apex by the above methods, he could cause the T-wave to appear or disappear.

The direct excitability of the heart muscle to external stimuli is also depressed during vagus stimulation. This effect is, however, not evi-

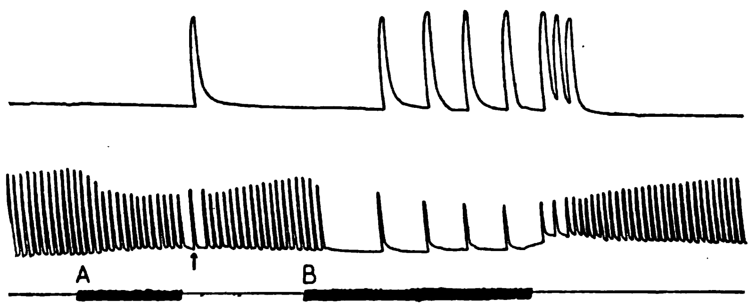


Fig. 67.—Tracing to show that vagus stimulation may facilitate transmission from auricles to ventricles. It shows the effect of right vagus stimulation on heart-block produced in the turtle by a clamp. Upper tracing records ventricle; lower tracing, auricles. Weak faradization of the right vagus nerve beginning at *A* affected the degree of block only at ↑, when a lengthened period between auricular contractions caused a single ventricular contraction. At *B* stronger faradization of the same nerve produced marked slowing of the auricles, in consequence of which the block disappeared and the ventricles contracted after each auricular contraction. Block reappeared when the rate again became rapid. Initial auricular rate = 36 per minute. (From Garrey.)

dent in the case of all hearts, but is seen in those of certain fishes (e. g., the eel).

The Mammalian Heart.—The action of the vagus on the mammalian heart may be investigated either by exposing the heart and connecting the auricles and ventricles with specially designed recording levers (myocardiograph), or if we desire to study the influence on the heart as a whole, by taking a blood-pressure tracing from one of the large arteries by means of a spring manometer. The results are in general similar to those observed on the frog or turtle heart, the main effects being developed on the auricle. Considerable differences are found in the effect on the heart as a whole in different animals, particularly with regard to the facility with which escapement occurs. In the dog when the vagus

is continuously stimulated, the heart is likely to remain inhibited for a long time, whereas in the cat the inhibition is very quickly broken into by escapement. If the tracing is taken directly from the heart, it will frequently be observed in the dog that, when the escapement occurs during vagus stimulation it is only the ventricle that is beating, the auricles still remaining inhibited.

If the stimulation of the vagus is discontinued after some time in an animal whose blood pressure is being recorded, the pressure will not only quickly recover, but will usually overshoot the normal level, mainly because of the asphyxia which has been produced during the period of inhibition. The asphyxia raises the hydrogen-ion concentration of the blood and this stimulates both the vasoconstrictor center and the heart action (page 168). The increased heart action is also partly owing to the fact that during vagus inhibition the beating power of the heart becomes improved (page 225).

As an outcome of recent work,[14] it has been shown that the right vagus nerve acts mainly on the sinoauricular node, and the left vagus on the auriculoventricular bundle. This is in agreement with the observations described above on the cold-blooded heart (page 217). Stimulation of the right vagus always causes slowing and weakening of both the auricular and the ventricular beats, but stimulation of the left vagus is sometimes observed to have but little influence on the auricular beat, although it may produce a condition of partial heart-block; or, if a clamp is applied to the auriculoventricular bundle so as to produce a partial heart-block, then during stimulation of the left vagus, the block may become complete. There is, however, a considerable overlapping of these influences, at least in the case of the left vagus, for this nerve also acts considerably on the ventricle, influencing perhaps not so much the rate as the force of the contraction. It has been found experimentally that, in order to demonstrate the specific action of the left vagus on the bundle, it is most suitable to study the relationship between auricular and ventricular beats when the auricle is beating rapidly as during the application of artificial (electrical) stimuli to it. Ordinarily the contraction produced by each stimulus passes into the ventricle, but during stimulation of the left vagus it is found that every contraction does not pass. These experiments raise the question as to what the influence of either nerve may be in blocking impulses from the auricles to the ventricles when auricular fibrillation is present. It might be expected that the left vagus would prove more effectual in this regard, but actually it has been found that both vagi have the same effect.

Tonic Vagus Action.—Impulses are constantly passing along the vagi to the heart. On account of this so-called tonic action, the heart rate

increases when the continuity of the vagus nerve is broken either by cutting or by freezing a portion of nerve (Fig. 26). The effect is usually inconspicuous when one nerve only is cut, but in most mammals it becomes quite marked when both are cut. Change in the heart rate produced by muscular effort is much more gradual in animals with marked vagus tone, such as hunting dogs, than in those with little vagus tone, as in domestic rabbits. The degree of vagus tone therefore bears a relationship to the staying power of the animal for prolonged muscular effort. It is usually ill developed in cold-blooded animals. It is quite marked in the case of man, as is evident on observing the heartbeat before and after giving a sufficient dose of atropine to paralyze the termination of the vagus in the heart.

The exact location of the nerve cells that form the center of discharging impulses along the vagus fibers to the heart can not be made out with certainty, but they are no doubt part of the great motor nucleus (ambiguus), from which arise the fibers not only of the vagus but of the glossopharyngeal nerve. The tone of this vagus center is almost without doubt dependent upon the constant transmission to it along the sensory or afferent fibers of impulses coming from various portions of the body. According to the strength or number of these impulses, the tone may be increased or diminished, thus altering the rate of the heart. It is possible of course that the tone can be maintained, independently of the afferent impulses, by the action on the center of chemical metabolic products or hormones produced in the cells or carried to them in the blood. We know at least that, like the respiratory center, that of the vagus is excitable by such hormones as the hydrogen-ion concentration of the blood. The tonicity of the vagus center is, however, mainly dependent upon the passage to it of afferent impulses, and as evidence for this conclusion may be cited the observation that, after section of most of the afferent nerves to the medulla (as by cutting the spinal cord high up in the cervical region), subsequent section of the two vagi does not produce anything like the usual degree of change in the heart rate.

The Afferent Vagus Impulses.—The afferent vagus impulses may come from practically any part of the body, having been first discovered by the simple experiment of tapping the abdomen of the frog with the handle of a scalpel, when slowing of the heart rate is observed. Cutting the vagi abolishes the reflex. Similar cardiac inhibition is produced by mechanical stimulation of the tail or gills of an eel. In mammals stimulation of the central end of any sensory nerve usually slows the heart, though sometimes the opposite effect occurs. The pulmonary branches of the vagus are particularly sensitive in producing reflex inhibition, and distinct results are usually obtained: by stimulation of the termina-

tions of the fifth nerve in the mucosa of the upper respiratory passages, as by inhaling ammonia vapor; by stimulation of the sensory nerve endings in the pharynx, as by swallowing; and of the mucosa of the larynx, as when a substance is "swallowed the wrong way." The sensory nerves of the abdominal viscera seem to be particularly active on the vagus center, as is seen in irritation of the sensory nerves of the stomach such as occurs in gastritis. Profound inhibition may also be caused by violent stimulation of the mesentery, as from a blow on the abdomen, or by irritation of the sensory nerves of the intestine, either mechanical or because of disease. Another interesting illustration of afferent vagus stimulation is obtained by pressure on the outer canthus of the eye. This oculomotor vagus reflex, as it is called, is very marked in some individuals.

Through which of these afferent paths it may be that the constant stimuli are transmitted to the vagus center to enable it to maintain its tone, can not be said, although it is very likely to be through the visceral nerves.

In considering the cause for an observed change in heart rate, we must of course bear in mind the possibility that the action may have occurred, not through the vagus center, but through the sympathetic center. Thus, when the heart becomes quicker, it may be owing either to diminution in the vagus tone or to an increase in the discharges along the sympathetic nerve from the augmentor center. That such reflex action through the augmentor center does occur under experimental conditions has been clearly shown; for example, if both vagus nerves are cut and the peripheral end of one of them stimulated moderately, so as to hold the heart at about its normal rate, the stimulation of certain sensory nerves may cause increase in the heart rate. Reflex sympathetic control of the heartbeat is however no doubt much less important than control through the vagus center. When it does exist it means that the actual rate of the heartbeat at any given moment must represent the algebraic sum of two opposing influences, with that of the vagus preponderating. The advantage of such a double innervation is that it insures prompter adjustment of the beat. If, for example, for any reason quickening of the heart rate is necessary, it is brought about most promptly if the vagus tone is diminished at the same moment that the sympathetic tone is increased. Such reciprocal action of antagonistic influences is the usual rule in the animal economy. Thus, when the knee joint flexes, it does so not merely because stimulating impulses are transmitted to the hamstring muscles, but also because at the same moment inhibiting impulses are transmitted to the extensor muscles (see page 814).

Several possibilities have to be kept in mind when we attempt to determine *the exciting cause for an observed change in the heart rate in man.* Thus, a slowing of the rate may be due to mechanical stimulation of the vagus trunk, as in pressure on the nerves by a tumor or aneurism in the neck. That such mechanical irritation may stimulate the vagus is easily demonstrated in many individuals by applying pressure to the vagus where it lies in the neck in front of the sixth cervical vertebra. Such pressure sometimes produces so profound an inhibition of the heart that temporary loss of consciousness occurs. It is often an unsafe experiment to perform.

Change in the heart rate in man may be caused by direct stimulation of the vagus center, as by the presence of a tumor or a blood clot in the medulla, or by the action on the center of some unusual hormone in the blood. A general increase in intracranial pressure also stimulates the vagus center. The slowing of the heart which occurs in asphyxia might be due either to the action of hormones (hydrogen-ion concentration) in the blood as the result of the asphyxia, or to the increased intracranial pressure. That the latter is the more important cause is shown by the fact that, if the rise in blood pressure is prevented by connecting an artery with a mercury valve,—that is, with a tube dipping into a cylinder of mercury to a depth corresponding to the normal blood pressure, so that when the pressure tends to rise the blood escapes,—the slowing of the heart is not observed. The excitability of the afferent vagus fibers in the lungs is greatly increased during the earlier stage of chloroform administration.

Finally it should be pointed out that, although we have no voluntary control of the activity of the vagus center, its activities are subject to great variation as a result of impulses transmitted from centers higher up in the cerebrospinal axis. It is by the operation on the vagus center of such impulses that changes in heart rate occur during emotional excitement, fright, etc. The increased heart rate in muscular exercise is probably dependent upon a number of causes, such as the irradiation of the motor impulses on to the cardiac centers (see page 412), the rise in temperature and changes in the hydrogen-ion concentration of the blood, etc.

Mechanism of Action of Vagus on the Heart.—Physiologists have naturally been curious as to the exact manner in which the vagus nerve brings about inhibition of heart action. Similar inhibition as a result of stimulation of efferent nerves exists in the case of the dilator fibers to the blood vessels (page 234) and the sympathetic nerve to the intestine (page 467). Inhibition of voluntary muscles can be produced only through the central nervous system by stimulation of afferent nerves

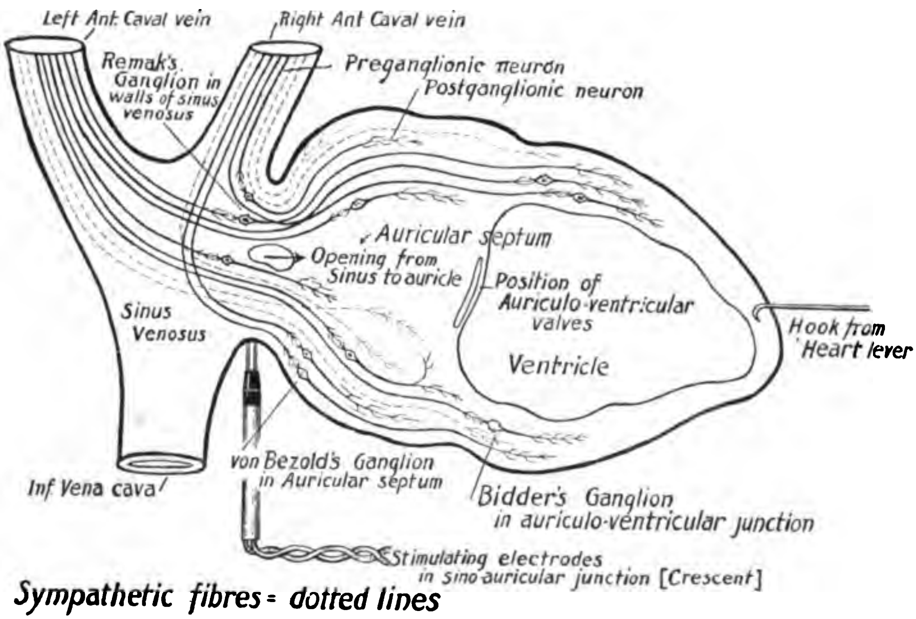


Fig. 68.—Diagram to show the innervation of the heart in the frog or turtle. The electrodes are represented as applied to the white crescentic line where they will stimulate some postganglionic fibers. (From Jackson.)

(page 814). It is not the nerve fibers themselves that are responsible for the inhibitory effect, for it has been found that if the peripheral end of a cut vagus nerve is connected with the central end of one of the anterior roots of the cervical portion of the spinal cord, the axons of the latter when they grow down into the vagus trunk during the regeneration which follows, stimulation of the regenerated fibers will still produce inhibition of the heart. The nature of the fibers can not therefore be the factor upon which the inhibiting influence of the vagus is dependent. This leaves the terminal apparatus of the vagus fibers in the heart as the structures in which the stimulus conveyed to them is rendered inhibitory in nature.

There has been considerable speculation as to what kind of change must be occurring in the heart in order to cause the inhibition, but practically nothing that is definite is known. One significant fact, however, is that the electrical current led off through nonpolarizable electrodes from two portions of the auricle one of which is injured, does not take the same direction when the vagus nerve is stimulated as that which it takes when the motor nerve of a similarly observed muscle is stimulated. A positive instead of a negative variation is observed. Now, since a negative variation is always accompanied by active chemical breakdown changes occurring in the muscle to supply its energy of contraction, it is assumed that the positive variation accompanying stimulation of the vagus must indicate that, instead of a katabolic process, a building up, or anabolic process, is being excited. This conclusion would fit in perfectly with the well-known fact that, after the heart has been held in standstill for some time by vagus stimulation, the beats are stronger after the inhibition has passed off than they were before. The vagus seems to have a conserving influence on the heart. During the inhibition produced by it energy material is apparently stored up in the heart, so that when the beat is reestablished it is stronger than before.

The Manner of Termination of the Vagus Fibers in the Heart.—This subject is of considerable pharmacologic and therefore therapeutic interest. In approaching the problem it must be remembered that the vagus fibers belong to the so-called cerebral autonomic system of nerves (see page 882). They are therefore fibers which have cell stations situated near their peripheral termination—cell stations, that is to say, in which ganglionic medullated fibers, by forming synapses around nerve cells, become connected with postganglionic nonmedullated fibers. The existence of ganglia in the heart, particularly of the frog, has been known for a long time. These ganglia are located at the sinoauricular junction, at the interauricular septum, and in the ventricle near the

auriculoventricular junction. The function of the ganglia is to serve as cell stations on the course of the vagus nerves. (Fig. 68.)

Nicotine is a drug which in certain concentrations, if applied locally to sympathetic ganglia, specifically paralyzes the synapses between the ends of the preganglionic fibers and the cells from which the postganglionic fibers arise. If this drug is applied in a 1 per cent solution to the heart, stimulation of the vagus trunk no longer produces inhibition, but if the stimulus is applied to a portion of the heart known as

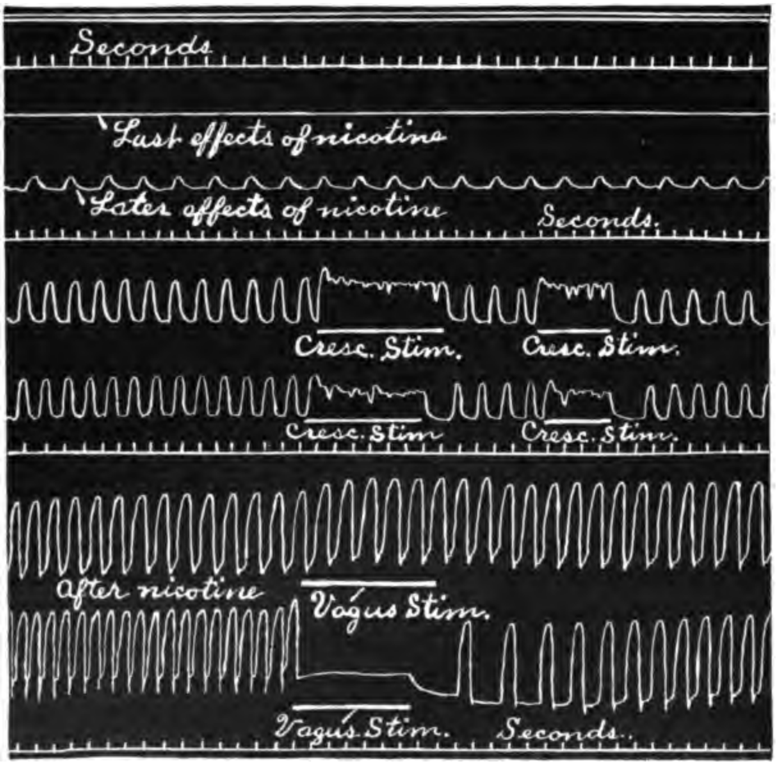


Fig. 69.—Frog heart tracing showing the action of nicotine. The vagus trunk was stimulated as indicated. In the normal (lower) tracing inhibition occurs but after nicotine (second tracing) no inhibition follows. Stimulation of the crescent in the next two lines still is followed by inhibition. The final effects of the drug are shown in the last two (upper) tracings. (From Jackson.)

the white crescentic line, inhibition still occurs, because at this point the postganglionic nerve fibers come near to the surface and therefore are stimulated (Fig. 69). On the other hand, atropine is a drug which paralyzes the postganglionic fibers, so that after its application to the heart inhibition can not be produced by stimulating either the vagus trunk or the white crescentic line. Pilocarpine and muscarine are drugs which have an action exactly opposite or antagonistic to that of atro-

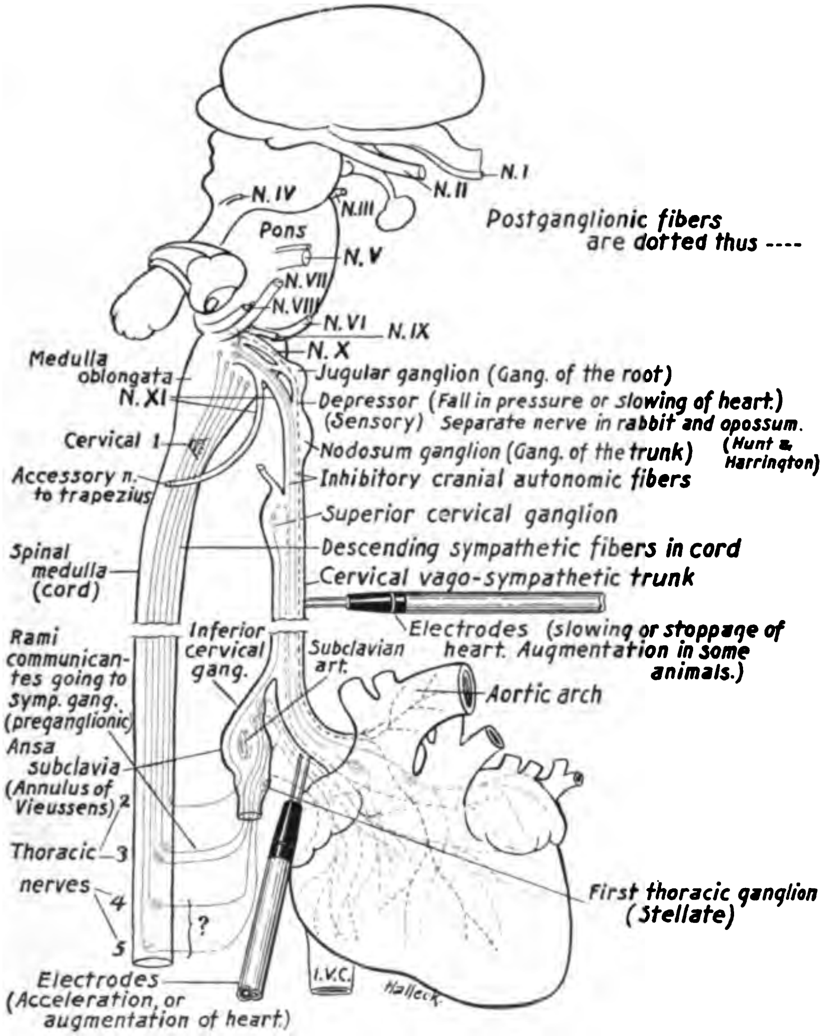


Fig. 70.—Schematic representation of the innervation of the heart of the mammal. The red continuous lines represent the sympathetic (accelerator) preganglionic fibers, and the broken red lines, their postganglionic fibers. The cell stations are in the inferior cervical and stellate ganglia, some extending up to the superior cervical ganglion. The green continuous lines represent the vagus preganglionic fibers, and the broken green lines, their postganglionic fibers. The cell stations in this case are located in the heart itself. It will be observed that electrodes applied to the so-called vagus low down in the neck may stimulate some sympathetic fibers. (From Jackson.)

pine; that is, they stimulate the postganglionic fibers and produce a slowing and possibly an enfeebling of the beat.

In the mammalian heart a large number of the fibers in the right vagus nerve proceed directly to the sinoauricular node, where it can be shown histologically that considerable masses of nervous tissue exist. On the other hand, the great majority of the fibers in the left vagus proceed to the auriculoventricular bundle, in which also nervous structures are abundant (page 184). As already indicated, the experimental results which follow stimulation of either nerve can be explained by the influence which the nerve exerts on the particular structure to which the majority of its fibers proceed. In brief, stimulation of the right vagus is likely to produce slowing and weakening of the beat, whereas stimulation of the left vagus is more likely to institute a condition of partial heart-block.

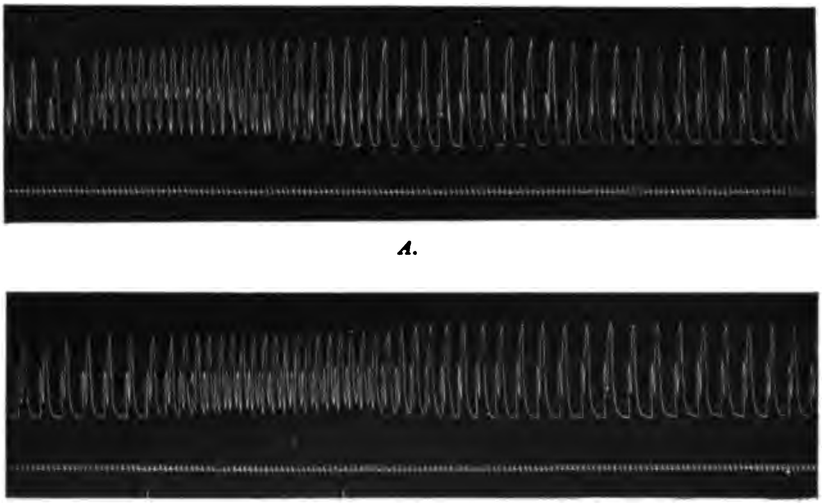
On account of the different results which may be obtained by stimulating the vagus, some authorities have assumed that the heart must contain *four kinds of fiber, more strictly, of vagus nerve endings,* one for each kind of influence which the vagus can develop. These four influences are, it will be remembered, on the strength, the rate and the propagation of the heartbeat, and the excitability of the cardiac muscle. It is, however, almost certainly unnecessary to make such an assumption, for the results can be explained as merely dependent upon different degrees of stimulation of the same kind of fiber and upon the exact part of the heart to which the fiber runs. Sometimes, for example, when the right vagus nerve is stimulated very feebly, there may be only a diminution in the force of the beats without any change in their rate, indicating that the effect has been upon the musculature of the auricular walls and not on the sinoauricular node. If the stimulus is increased a little, then both an enfeebling and a slowing of beat occur, indicating that the stimulus has now passed both to the auricular musculature directly and to the sinoauricular node.

The Sympathetic Control

The effect of the sympathetic nerve on the heart may be described as being exactly opposite to that of the vagus. The pathway along which the fibers of this nerve travel to the heart is more or less a devious one. They arise in the mammal from nerve cells in the gray matter in the upper thoracic portion of the spinal cord. The fibers leave by the corresponding spinal roots and pass by the white rami communicantes into the sympathetic chain, up which they travel to the stellate and inferior cervical ganglia. Around the nerve cells of the stellate ganglion the fibers end by synapsis, and the axons of the cells are then continued on

as postganglionic fibers, proceeding to the heart through branches coming off from the stellate ganglion itself, or from the ansa subclavii or inferior cervical ganglion. (Fig. 70). In cold-blooded animals, such as the frog, the sympathetic fibers run up to the upper end of the cervical sympathetic and join the vagus immediately after it leaves the cranial cavity. They then proceed along with this nerve—forming the vago-sympathetic—to the heart. The effect of stimulation is shown in Fig. 71.

The sympathetic nerve differs from the vagus in that a much longer latent period elapses before its influence becomes effective, and this persists for a much longer period after the stimulus is withdrawn. If the vagus



A.

B.

Fig. 71.—Tracings showing the effects on the heartbeat of the frog resulting from stimulation of the sympathetic nerves prior to their union with the vagus nerve. (From Brodie.)

and sympathetic are stimulated at the same time, as by exciting the vago-sympathetic in the frog, the first effect observed is that of the vagus usually followed, after removal of the stimulus, by the sympathetic effect. If the stimulus is maintained for a long time, so that the vagus becomes fatigued, escapement will occur earlier than with pure vagus stimulation, and augmentation may become apparent. The sympathetic influence is, however, never so strong as that of the vagus. The two nerves are therefore not antagonistic in the sense that the one neutralizes the effect of the other; but when both are stimulated, the heart responds first to the vagus and later to the sympathetic.

CHAPTER XXVI

THE CONTROL OF THE CIRCULATION (Cont'd)

THE NERVE CONTROL OF THE PERIPHERAL RESISTANCE

As already explained, the nerve control of the peripheral resistance takes place through the action of vasoconstrictor and vasodilator nerve fibers on the musculature of the arteriole walls. The vasoconstrictor impulses like those in the vagus of the heart are tonic, so that when a nerve containing such fibers is cut, the corresponding blood vessels undergo dilatation (see page 135), and when their peripheral ends are stimulated artificially, constriction occurs. On the other hand, the vasodilator impulses do not appear, at least under ordinary circumstances, to be tonic, so that the cutting of such fibers does not cause vasoconstriction; their stimulation, however, causes marked dilatation. Vasomotor fibers are contained in most of the efferent (motor) nerve trunks, and to detect their presence the nerve must be either cut or stimulated and the condition of the blood vessels of the innervated area observed.

Methods for the Detection of Constriction or Dilatation

Several methods, varying with the exact area under observation, can be used *for the detection of vasoconstriction or dilatation*. In many cases *visual inspection* is sufficient, as in the well-known experiment of Claude Bernard on the blood vessels in the ear of the rabbit (see Fig. 106). When this is held with a light behind it, and the cervical sympathetic of the corresponding side is cut, marked dilatation will become evident and vessels will spring into view where previously there were none visible. Visual inspection is usually also a satisfactory method of demonstrating vasodilatation or constriction in exposed glands, in mucous passages and in the vessels of the skin.

Another comparatively simple method is the observation of *the temperature of the part*, this being particularly useful when the vascular area is one situated in the peripheral part of the body, such as the hand or foot (see page 209). When dilatation occurs the temperature of the part rises, because the warmer blood from the viscera flows with greater freedom through the peripheral regions, where it is cooled off by radiation. When a thermometer is placed between the toes of a dog or cat, a

distinct rise in temperature will be observed when the sciatic nerve of the corresponding limb is cut. The application of this principle in determining the mass movement of blood by the amount of heat given off from the hands or feet has already been explained.

Other methods depend upon observation of the *outflow of blood from the veins of the part.* A simple application of this method can be used in the case of the ear of the rabbit. If the tip of the ear is cut off, bleeding under ordinary circumstances is only very slight, but if the cervical sympathetic is cut, it becomes quite marked, slowing down again or even stopping entirely when the peripheral end of the nerve is stimulated. By making measurements of the volume of the outflow of blood from a vein by this method, the extent of constriction or dilatation can

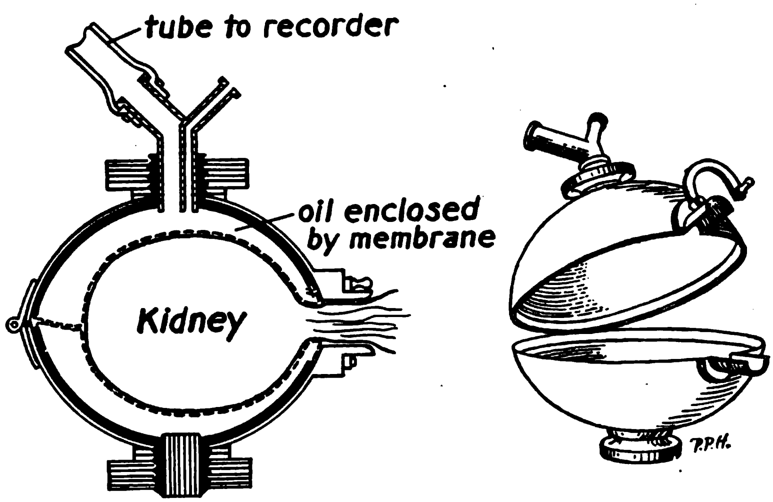


Fig. 72.—Roy's kidney oncometer. (From Jackson.)

be followed quantitatively. Vasodilatation also causes changes in the *character of the venous flow,* the usually continuous flow becoming pulsatile and the color of the blood brightening. Comparison of the *pressures in the arteries and the veins of* a part is also often of value in the detection of changes in the caliber of the blood vessels, for, of course, the greater the difference in pressure between the two manometers, the greater must be the resistance offered to the flow.

For experimental purposes, however, the standard method is that known as the *plethysmographic.* For this purpose the organ or tissue is enclosed in a so-called plethysmograph or volume recorder, the principle of which will be clearly seen by consultation of the accompanying diagram of one adapted for the kidney (Fig. 72). Any increase detected by this means in the volume of the part must be due either to

an increase in blood flowing into the vessels because of increased heart action or to a local vasodilatation; and vice versa, when shrinkage occurs. We can not tell from the volume tracing itself which of these changes is really responsible for the observed alteration, but we can do so by simultaneously observing the mean arterial blood pressure. If this falls when the volume decreases, it means that the volume of blood flowing to the part must have become diminished. If, on the other hand, the blood pressure remains constant or rises while the volume decreases, it means that the blood vessels have locally constricted.

Methods for the Detection of Vasomotor Fibers in Nerve Trunks

If we wish to find out through which nerve trunks a given vascular area is supplied with vasoconstrictor or vasodilator impulses, we should proceed by the use of one of the above described methods to observe the effect produced on the vessels by cutting the nerve and then by stimulating the peripheral end of the cut nerve. As a result of such observations it has been found that the vasomotor fibers are frequently distributed so that those having a vasoconstricting action are collected mainly in one nerve trunk and those having a dilating action in another; in some nerve trunks, however, the relative numbers of the opposing fibers are about equal. Nerves containing a great preponderance of vasoconstrictor fibers are the great splanchnic and the cervical sympathetic; and those containing a great preponderance of vasodilator are the chorda tympani nerve to the submaxillary gland and the nervi erigentes to the external genitalia.

It must be clearly understood that, although one kind of vasomotor fiber may preponderate in one of these nerves, yet the opposite kind is also present. In the cervical sympathetic, for example, some vasodilator fibers extending to the blood vessels of the mucous membrane of the nose and cheeks can readily be demonstrated, as shown by the flushing of these parts when the peripheral end of the nerve is stimulated; and similarly, even in the great splanchnic nerve itself, vasodilator fibers supplying the suprarenal capsule can quite readily be made out. When the vasoconstrictor fibers greatly preponderate over the vasodilator, the effect of the latter may be demonstrated by taking advantage of the fact that *ergotoxine* paralyzes the vasoconstrictor but not the vasodilator fibers, so that after its administration stimulation of the great splanchnic nerve gives rise to a vasodilatation instead of a vasoconstriction. The presence of vasoconstrictor fibers in the so-called vasodilator nerves (chorda tympani and nervi erigentes) has not however, been demonstrated.

A good example of a nerve trunk containing about an equal admix-

ture of both kinds of vasomotor fibers is the sciatic. If the hind limb of a dog is placed in a plethysmograph and simultaneously a record of the mean arterial blood pressure taken, it will be found on cutting the sciatic nerve that the volume of the limb increases, whereas the blood pressure remains practically constant. Before placing the limb in the plethysmograph, the muscles must of course be paralyzed by means of curare; otherwise muscular contractions would confuse the result. If the peripheral end of the cut nerve is now stimulated, vasoconstriction will readily be observed. So far, then, the results demonstrate the presence of vasoconstrictor nerve fibers alone.

To demonstrate the presence of vasodilators a different procedure is necessary. This is based on the following facts: (1) The vasodilator nerve fibers degenerate more slowly than the vasoconstrictor; (2) they are less depressed in their excitability by cooling the nerve; and (3) they are more sensitive to weak slow faradic stimulation than the vasoconstrictor fibers. Accordingly, if we cut the sciatic nerve two or three days before the actual experiment, and then, while observing the volume of the limb, proceed to stimulate the half-degenerated nerve with feeble electric stimuli of slow frequency we shall usually observe a dilatation of the limb instead of constriction; and even if we cool a stretch of a freshly cut nerve before applying the stimulus, the same result will often be obtained.

The Origin of Vasomotor Nerve Fibers

Having seen how the presence of vasomotor fibers may be detected in peripheral nerves, we must now proceed to trace them back to their origin from the central nervous system. The method for doing this consists, in general, in observing the effect on the blood vessels produced by cutting or stimulating the various nerve roots through which the fibers might pass on their way to the nerve trunks. As a result of such observations it has been found that all of the *vasoconstrictor fibers* emanate from the spinal cord in the region between the level of the second thoracic and that of the second or third lumbar spinal roots, but from nowhere else in the cerebrospinal axis. Section of the spinal cord below the level of the second lumbar spinal roots produces no change in the volume of the hind limb, provided the muscles be thoroughly curarized, nor does stimulation of the lower end of the cut spinal cord have any effect. If the last two thoracic or the first two lumbar spinal roots are stimulated, however, evidence of vasoconstriction will be obtained.

The restriction of the origin of vasoconstrictor fibers to the above-mentioned regions of the spinal cord indicates that in proceeding to the mixed nerve trunks they must travel along special nerve paths.

These are provided by the sympathetic chain and its branches (Fig. 228). The vasoconstrictor fibers in the anterior spinal roots leave the latter by way of the corresponding white rami communicantes, and pass into the neighboring sympathetic chain, along which they either ascend or descend, according to their ultimate destination. In their course they come into contact with the sympathetic ganglia, through one or two of which they may pass without any change, but ultimately each fiber arrives at some ganglion, in which it terminates by forming a synapsis around one of the ganglionic nerve cells. The axon of this nerve cell then continues the course by the nearest gray ramus communicans back to the spinal nerve beyond the union of its anterior and posterior roots. Up to the point where the fiber forms a synapsis with a ganglionic nerve cell, it is medullated and is known as the preganglionic fiber. Beyond the nerve cell, it is nonmedullated and is known as postganglionic (page 877).

The exact ganglion in which a given vasoconstrictor fiber becomes connected with a nerve cell can be determined by the nicotine method of Langley. Local application to the ganglion of a weak solution of this drug (1 per cent) paralyzes the synaptic connection, so that a stimulus applied to the preganglionic fiber no longer produces its effect. Suppose, for example, that a vasoconstrictor fiber has been found by the stimulation method to travel through several ganglia, and we wish to determine in which of these the synapsis occurs: we can do so by applying the stimulus at a point central to the ganglia after painting each of them in turn with the nicotine solution. If the application of the drug to a given ganglion is found to cause no alteration in the effect produced by stimulation, then we know that there can not be any synaptic connection in that ganglion, and we proceed in the same way till we have located the ganglion in which synapsis occurs. It is important to remember that the postganglionic vasoconstrictor fibers in a gray ramus communicans do not come from the preganglionic fibers of the corresponding spinal root, but from fibers coming through white rami at a higher or a lower level.

The above description applies to the vasoconstrictor fibers proceeding to the vessels of the anterior and posterior extremities. those for the former arising (in the dog) from about the fourth thoracic to the tenth; and those for the latter, from the lowest thoracic and the first three lumbar nerve roots. The cell station for the fibers to the fore limbs is in the stellate ganglion, and for the hind limbs in the last two lumbar and first two sacral ganglia of the abdominal sympathetic chain.

The vasoconstrictor fibers to the vessels of the head and neck run a somewhat different course, there being no convenient cerebrospinal nerve along which the postganglionic fibers may run. The fibers to the blood vessels of the head leave the cord by the second to the fourth or fifth thoracic roots and pass by the corresponding white rami communicantes into the sympathetic chain, up which they run, passing through the stellate ganglion, the ansa subclavii, and the inferior cervical ganglion, then ascending in the cervical sympathetic to the superior cervical ganglion, where their cell station exists. The postganglionic fibers on leaving this ganglion travel to their destination mainly along the outer walls of the blood vessels.

The vasoconstrictors to the abdominal viscera are carried by the splanchnic nerves, the fibers of which come off from the lower seven thoracic and the uppermost lumbar

roots. The thoracic fibers pass down the sympathetic chain, which they leave by the great splanchnic nerves. The lumbar fibers form the lesser or abdominal splanchnic nerves. As preganglionic fibers, therefore, these fibers are carried by the greater and lesser splanchnic nerves into the abdomen, where the former comes into close relationship with the suprarenal glands, giving off a branch to the suprarenal ganglion. The main course of the nerve is continued on to the solar plexus, in the various ganglia of which most of the preganglionic fibers end by synapsis, the postganglionic fibers then proceeding along the blood vessels to the vessels of the abdominal viscera. (See also page 879).

Vasodilator fibers have a more varied origin than vasoconstrictor, and they run an entirely different course. Vasodilator impulses may be transmitted by fibers arising from practically any level of the cerebrospinal axis, not only by the motor roots, but by the sensory as well. Thus, they pass out of the spinal cord in the posterior sacral roots to enter the nerves of the hind limbs, as has been demonstrated by observing an increase in the volume of the curarized limb during electrical stimulation of the exposed rootlets. The apparent inconsistency of these observations with the well-known law concerning the direction of the impulses contained in the posterior spinal roots is explained by assuming that the dilator impulses are transmitted along the ordinary sensory fibers, since there are no efferent fibers in these roots. They are impulses which go against the ordinary stream (antidromic). In support of this explanation it is of importance to note that at their termination near the skin many sensory fibers split into several branches, some of which run to blood vessels, and others to receptor organs (page 797). Stimulation of the latter branches may cause dilatation of the local blood vessels nearby, indicating that impulses must be transmitted up to the point at which the branching occurs and then down the vascular branch, this result being obtained even after the main trunk of the nerve has been cut above the division.

For the blood vessels of the anterior extremity, the vasodilator impulses are similarly transmitted through the posterior spinal roots of the lower cervical region of the spinal cord. The vasodilator fibers to the abdominal viscera are transmitted with the splanchnic nerves, but they may also be derived from the posterior spinal roots, for it has been found that stimulation of posterior roots in the splanchnic area causes dilatation in the intestine (Bayliss). Vasodilator fibers are also contained in the cranial nerves, particularly the seventh and the ninth, being distributed in the former nerve to the anterior portion of the tongue and the salivary glands, and in the latter to the posterior portion of the tongue and the mucous membrane of the floor of the mouth. The vasodilator fibers for the mucous membrane of the inside of the cheeks and nares have their course in the cervical sympathetic, being distributed to the buccofacial region in the branches of the fifth cranial nerve.

There is evidence to show that the vasodilator fibers, like the vasoconstrictor, become connected by synapsis with nerve cells somewhere in their course. In the case of the vasodilator fibers in the chorda tympani and nervi erigentes, such cell stations have been clearly demonstrated in the hilus of the submaxillary gland in the former nerve

and in the hypogastric plexus situated on the neck of the bladder in the latter. It is therefore commonly assumed that, although not recognizable by histologic methods, such terminal cell stations must also exist in close association with all blood vessels to which the vasodilator fibers run. Whether or not such peripheral cell stations exist, there is a marked difference between the course of vasodilator and of vasoconstrictor fibers.

The Vasomotor Nerve Centers

Our next problem is to trace these fibers farther into the central nervous system, and find the location and study the characteristics of the nerve centers from which they are derived. We must postulate the existence of both vasoconstrictor and vasodilator centers, but since there is no adequate evidence at the present time which enables us to locate the latter, we must confine our attention to the vasoconstrictor centers. These exist at two levels in the cerebrospinal axis: (1) in the gray matter of the spinal cord, and (2) in the gray matter of the medulla oblongata.

The spinal, or as they are often called, the subsidiary vasoconstrictor centers, are represented by certain cells of the lateral horn of gray matter in the thoracic portion of the spinal cord, from which the preganglionic vasoconstrictor fibers above described are derived. The exact location of the nerve cells composing the chief centers in the medulla has not as yet been definitely made out; they undoubtedly lie near those of the vagus center (see Ranson). The axons of the medullary cells descend in the lateral columns of the spinal cord to end by synapses around the cells of the subsidiary vasoconstrictor center in the lateral horns.

The experimental evidence which indicates the existence of chief and subsidiary centers is quite definite. Thus, if the spinal cord is cut at the lower cervical region (below the phrenic nuclei, so as not to interfere with the movements of the diaphragm), the arterial blood pressure falls profoundly, because the pathway connecting the two centers is broken. After several days, however, the blood pressure will gradually rise again. If after this has occurred, the spinal cord is destroyed by pushing a wire down the vertebral canal, the arterial blood pressure will again fall, indicating that the vascular tone which had been reacquired after section of the pathway between the main and the subsidiary centers must have been brought about by the development in the subsidiary centers of an independent power of reflex tonic action. This experiment therefore demonstrates that in the intact animal the subsidiary centers do not by themselves discharge tonic impulses. In other words, the subsidiary centers ordinarily serve merely as transfer stations for the tonic impulses coming from the chief center, but when these impulses no longer

arrive, then a hitherto dormant power of tonic activity becomes developed in the subsidiary centers.

Independent Tonicity of Blood Vessels

Even after complete disconnection of the spinal cord from the blood vessels, as by cutting of the splanchnic nerve to the abdomen or ablation of that portion of the lower spinal cord from which the fibers to the hind limb arise, the disconnected blood vessels, although at first completely dilated, may later reacquire an independent tone of their own, indicating therefore, that they must possess some neuromuscular mechanism which can act independently of the nerve centers, and which may be stimulated to activity by the presence of hormones in the blood. The hormone was at one time thought to be epinephrine (see page 745).

Epinephrine control is indicated in the effect produced upon arterial blood pressure by stimulation of the great splanchnic nerve. Careful analysis of the curve, shown in Fig. 29, shows that the rise is both immediate and delayed; that is, the curve mounts immediately, then flattens out a little, and then assumes a further rise. This delayed response seems to depend upon the excretion of epinephrine into the blood, for it does not occur when the suprarenal veins are occluded, and is much delayed by temporarily clamping the suprarenal veins on the same side as that on which the splanchnic nerve is stimulated. It has been stated by certain observers that, after occlusion of the adrenal veins, there is a downward tendency of the blood pressure, which however develops with extreme slowness; and that a distinct elevation of blood pressure follows the removal of a clamp temporarily placed on the adrenal veins. This rise is pronounced if the splanchnic nerve is stimulated during the occlusion of the veins. It must of course be understood that the immediate rise in blood pressure following splanchnic stimulation is caused by vasoconstriction in the splanchnic area itself, as is evidenced by the fact that it does not occur, or is only very faint, when the abdominal blood vessels are ligated prior to the stimulation of the splanchnic nerve. Even after ligation of the adrenal veins and of the blood vessels of the splanchnic area, stimulation of the splanchnic nerve may still cause a slight rise in arterial blood pressure, possibly because some fibers may run from the splanchnic to vascular areas not situated within the realm of the splanchnic nerve—for example, the blood vessels of the lumbar muscles.

CHAPTER XXVII

THE CONTROL OF THE CIRCULATION (Cont'd)

CONTROL OF THE VASOMOTOR CENTER

The activities of the vasomotor center are controlled partly by hormones and partly by afferent impulses.

The Hormone Control

As with the respiratory center, the chief hormone is the hydrogen-ion concentration of the blood. When this is increased, as in asphyxia, the vasoconstrictor part of the vasomotor center becomes stimulated, so that the blood vessels are constricted and the blood pressure rises. Taking, as our criterion of hydrogen-ion concentration, the tension of the carbon dioxide in the blood (see page 354), we may proceed to investigate the relationship by observing the blood pressure during changes in the carbon-dioxide tension brought about by causing the animal to breathe atmospheres containing known percentages of the gas (Mathison[15]). Thus, if a decerebrate cat is made to respire an atmosphere containing 5 per cent or more of carbon dioxide, an immediate rise occurs in the arterial blood pressure. That the inhaled carbon dioxide acts by raising the hydrogen-ion concentration of the blood is indicated by the fact that a similar rise in blood pressure can be obtained by intravenous injection of a weak solution of lactic acid (2 c.c. N/15) in a decerebrate cat.

Instead of injecting the lactic acid, we may cause it to be produced in the muscles of the animal itself by greatly diminishing their oxygen supply. When a decerebrate cat, for example, is made to breathe an atmosphere of almost pure nitrogen, there is, after a latent period of about 30 seconds, a sudden rise in arterial pressure. The existence of this latent period in the latter case, as compared with its absence when carbon dioxide is inspired, is owing to the time taken for lactic acid to be produced in the muscles on account of the oxygen deprivation. It is important to note in the above experiment that decerebrate animals are employed so as to avoid the necessity of using anesthesia, under which the results are much less definite. The fact that oxygen deprivation causes excitation of the vasoconstrictor center has been known for

some time, but the explanation that has usually been given has been that it is due to a direct effect of oxygen want on the center.

The sensitivity of the medullary center towards the hydrogen-ion is many times greater than that of the subsidiary centers in the spinal cord. If an animal is kept alive by artificial respiration for some time after cutting the cervical spinal cord, the subsidiary vasomotor centers will, as we have seen, gradually acquire a tonic action, and the lowered blood pressure will gradually rise again. If, when this has been attained, the animal is made to breathe an atmosphere rich in carbon dioxide, a sudden rise in blood pressure will occur, but to produce it a very much greater percentage of this gas must be inspired than when the pathway between the chief and subsidiary centers is intact. Whereas 5 per cent carbon dioxide is sufficient to cause a rise of pressure in an animal having its chief vasomotor center, it takes 25 per cent and upward to produce a like effect on a spinal animal; and similarly, although 2 c.c. of N/15 lactic acid will stimulate the chief vasomotor center, it takes 5 c.c. of N/2 to excite the spinal-cord centers.

The Nerve Control

However important hormones may be in maintaining a tonic stimulation of the center, the more sudden changes in activity are mainly brought about by afferent nerve impulses. The afferent impulses are of two classes: (1) those causing a rise in blood pressure, called *pressor,* and (2) those causing a fall in blood pressure, called *depressor.* The effect produced on the arterial blood pressure by stimulation of either pressor or depressor fibers is usually more or less evanescent, especially in the case of the depressor fibers; and when the change following stimulation of the nerve passes off, the blood pressure always returns to its former level. This indicates that the afferent impulses do not affect the tonic control which the vasomotor center exercises on the blood vessels. It has, therefore, been assumed by Porter[16] that there are really two kinds of vasomotor centers: one concerned merely in the bringing about of temporary reflex changes, the other concerned in the maintenance of the vascular tone. It may be that the activities of the former are primarily dependent upon afferent impulses, and the latter, upon hormones. Justification for this view has been found in observations made on the effects of stimulation of pressor and depressor fibers in animals under the influence of curare or alcohol. With the former drug, stimulation of a nerve containing a preponderance of pressor or depressor fibers produces double its usual effect, but the mean level of the blood pressure apart from this effect remains unchanged. With the latter drug (alcohol), on the other hand, the reflex response entirely

disappears, although it immediately reappears when the alcohol effect has passed off, and there is no evidence of a change in tone. The tonic and the reflex mechanisms of the vasomotor center can not therefore be identical.

At the present stage of our knowledge, it is only possible for us to study the effect of stimulation of pressor and depressor fibers on the vasoreflex center. Such fibers are contained in practically every sensory nerve of the body, and it would appear that a fairly equal mixture of both kinds of fiber exists in most of these nerves.

Pressor and Depressor Impulses.—Depressor impulses are alone present in the *cardiac depressor nerve.* Sometimes as in the rabbit, this exists as an independent nerve trunk, originating by two branches, one from the superior laryngeal, the other from the vagus, and descending close to

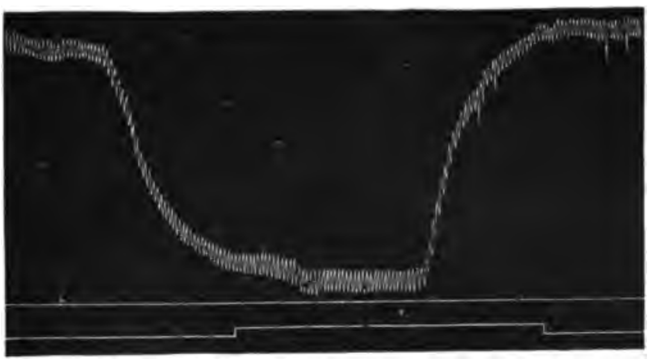


Fig. 73.—Fall of blood pressure from excitation of the depressor nerve. The drum was stopped in the middle of the curve and the excitation maintained for seventeen minutes. The line of zero pressure should be 30 mm. lower than here shown. (From Bayliss.)

the vagus trunk, to end around the arch of the aorta. In other animals the depressor is bound up with the vagus trunk from which it can sometimes be separated by careful dissection. The first prerequisite in investigating the cause of the changes produced by stimulation of these nerves is the elimination of any chance of an alteration in heartbeat as a result of simultaneous stimulation of afferent vagus fibers. This may be done either by cutting both vagi or by administering atropine.

Stimulation of the central end of the cardiac depressor nerve in such an animal causes an immediate fall in blood pressure, accompanied by an increase in volume which can be demonstrated either in the hind limb or in one of the abdominal viscera—evidence of general vasodilatation (Fig. 73).

When the central end of a sensory nerve, such as the sciatic, is acted on by a stimulus of moderate strength, it will usually be found that the arterial blood pressure rises and that the volume of the limb or of some

abdominal viscus becomes diminished—evidence of general vasoconstriction. But when the sensory nerve is stimulated with extremely weak faradic shocks, an entirely different result is likely to be obtained; namely, a fall of blood pressure and an increase in volume of the limb or viscus, indicating that in this manner we have stimulated depressor fibers. By careful experimentation with quantitatively graduated electrical stimuli, it has been found by Martin and others[17] that on stimulating an afferent nerve with weak shocks, a fall in blood pressure is the first effect to be observed, and that this becomes more and more marked as the strength of the stimuli is increased, until a certain optimum is reached, after which the fall in blood pressure becomes less evident. When a certain strength of stimulation is exceeded, a rise instead of a fall occurs. After this point additional increase in stimulation causes more and more marked elevation of blood pressure through a very long range of stimuli.

Stimulation of two afferent nerves at the same time usually produces a greater reflex vasomotor change than the stimulation with an equivalent strength of current of either nerve alone. That is to say, the effect produced by stimulating the central end of both sciatics simultaneously will be greater than that produced by stimulating either alone with double the strength of stimulus.

As has been stated above, the reflex change in blood pressure is often quite transitory in nature, although the stimulation of the pressor nerve is maintained. When this decline has occurred, the pressor reaction can often be renewed by shifting the stimulation to a second nerve. These facts concerning the greater efficacy of combined stimulation of several nerves are of considerable importance in connection with the general question of reflex changes in blood pressure. For instance, many of the pressor fibers found in the sciatic nerve are connected with the receptors that mediate the sensations of the skin. When these receptors are stimulated, as by heat or cold, reflex changes in blood pressure occur (pressor reaction), (Fig. 74), and it is important to remember that localized stimulation of the skin is less efficient in bringing about such vascular changes than stimulation applied over large areas, even when the local stimulus is intense and the general stimulus mild in character. Jumping into a moderately cold bath will cause a much greater rise in arterial blood pressure than plunging the hand into ice cold water.

Mechanism of Action of Pressor and Depressor Impulses.—When we consider the exact mechanism by which these afferent impulses operate, we have to bear in mind four possibilities: the reflex fall produced by stimulation of a depressor afferent fiber may be due either to a stimulation of the vasodilator part of the center or to an inhibition of the tone

cf the vasoconstrictor part; and, conversely, a rise in arterial pressure caused by vasoconstriction may be dependent either on a stimulation of the vasoconstrictor part of the center or on an inhibition of the tone of the vasodilator part. All of these changes have, as a matter of fact, been shown to occur, at least under certain conditions, although the evidence

Fig. 74.—The effect of strong stimulation (heat) of the skin of the foot on the arterial blood pressure and respiratory movements. Upper tracing, thoracic movement; lower tracing, arterial blood pressure.

for the inhibition of dilator tone is as yet a little uncertain (see Fig. 75).

Without going into the subject in detail, we may nevertheless take as an example of the methods by which the information has been obtained, the experiment performed by Bayliss,[18] showing that the vasodilation which results from stimulation of the depressor nerve is owing partly to removal of vasoconstrictor tone and partly to vasodilator

stimulation. The volume of the hind limb of a curarized and vagotomized rabbit increases when the central end of the cardiac depressor nerve is stimulated. In order to determine whether this dilatation is due solely to the removal of vasoconstrictor tone, the above experiment was repeated on a rabbit in which the sympathetic chain had been cut below the level of the second lumbar spinal roots. By such an operation all the vasoconstrictor fibers to the vessels of the hind limb are severed, but the vasodilator fibers, since they emanate through the sacral sensory roots, are left intact. It was nevertheless found on stimulating the depressor nerve that dilatation of the hind limb still occurred, thus indicating

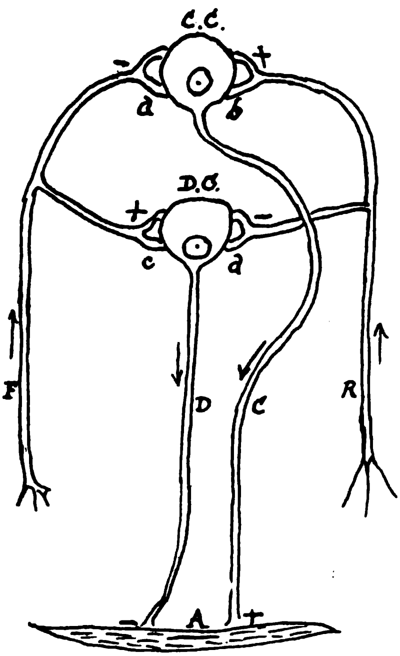


Fig. 75.—Diagram showing the probable arrangements of the vasomotor reflexes.

A. Muscle of arteriole.
D. Vasodilator nerve fiber terminating on *A* and inhibiting its natural tonus, as indicated by − sign.
D. Vasoconstrictor fiber also ending in *A*, but exciting it (+). These two kinds of fiber arise from the dilator center (*DC*) and the constrictor center (*CC*) respectively.
F. Afferent depressor fiber, dividing into two branches, one of which (−) inhibits the constrictor center, while the other (+) excites the dilator center causing dilatation of the arteriole and fall of blood pressure.
R. Pressor fiber exciting *CC* and inhibiting *DC*, and therefore causing vasoconstriction and rise of blood pressure.
a, *b*, *c*, and *d* represent the synapses of the pressor and depressor branches with the efferent neurons. (From Bayliss.)

that stimulation through vasodilator fibers must have taken place. Conversely, in another experiment, instead of the sympathetic chain, the spinal cord was cut below the level of the second lumbar segment, thus

severing the dilator but not the constrictor path, and again depressor stimulation caused the volume of the limb to increase, indicating that an inhibition of constrictor tone must have occurred.

Reciprocal Innervation of Vascular Areas

It must not be imagined that changes in the caliber of the blood vessels occurring in one vascular area are necessarily occurring all over the body. On the contrary, a most important reciprocal relationship exists in the blood supply to different parts. *After food is taken,* for example, more blood is required by the digestive organs than when they are at rest, and this is insured by dilatation of their own vessels along with reciprocal constriction of those of other parts of the body. On account of the relatively great capacity of the abdominal vessels, their dilatation during digestive activity is usually greater than the reciprocal constriction of the other vessels, so that the diastolic blood pressure falls, necessitating a more powerful cardiac discharge in order to maintain the mean pressure. After taking food, the systolic pressure does not as a rule fall so much as the diastolic, if it falls at all; and the pressure pulse therefore becomes greater and causes a greater live load to be applied to the vessels with each heartbeat. During the sudden strain that is thrown on them, weakened arteries may give way, especially in the brain.

Another example of reciprocal action of the vascular system is seen in *muscular exercise.* The vessels of the active muscles dilate, while those elsewhere constrict. The local dilatation in this case is, however, not entirely at least a nervous phenomenon, being caused in fact, as we shall see, by hormone action on account of the local increase in hydrogen-ion concentration (see page 414). There can be little doubt that *local irritants* to the surface of the body, such as hot applications, liniments, etc., act in the same way; they cause local dilatation of the superficial and perhaps of the immediately underlying vessels and constriction of those elsewhere in the body. Application of cold to local areas of skin similarly causes local constriction accompanied by reciprocal dilatation elsewhere. This action of cold is very marked in some parts of the body, such as the hands, where by Stewart's method (page 283) it can be shown, not only that the bloodflow of the hand to which the cold is applied is greatly curtailed, but also that of the opposite side.

Experimental demonstration of reciprocal vascular innervation is furnished by numerous experiments. If the central end of the great auricular nerve of the ear is stimulated in a rabbit, a blanching of the vessels of the ear occurs at the same time as a rise in arterial blood pres-

sūre (Lovén reflex). Similarly when the central end of one of the sensory roots of the leg of a dog is stimulated, there is a rise in arterial blood pressure and an increase in the volume of the limb.

THE INFLUENCE OF GRAVITY ON THE CIRCULATION

If the arterial blood pressure is measured in the arm and leg in a man standing erect, a difference corresponding to the hydrostatic effect of gravity will be found between the two readings. In comparison with the high pressure normally existing in the arteries, this difference is, however, of little significance. On the other hand, in the veins, where the average pressure is low, gravity would cause serious embarrassment to the circulation of blood were it not for the valves and the forces which move the blood beyond them (page 214).

In erect animals the part of the circulation in which blood might stagnate as a result of gravity is the splanchnic area. Were such stagnation to occur, the blood would not be returned to the right heart, so that the arteries would not receive sufficient blood to maintain an adequate circulation, particularly in the vessels of the brain.

Simple experiments devised by Leonard Hill[19, 25] illustrate these principles. When a snake, for example, is pinned out on a long piece of wood and an opening made opposite the heart, this organ can be seen to fill adequately with blood as long as the animal is maintained in the horizontal position. When placed vertically, however, the heart becomes bloodless. If now the tail end of the animal is placed in a cylinder of water so as to overcome the effect of gravity, the heart will be seen to fill again with blood. Evidently in such an animal there is no mechanism to compensate for gravity.

If a domestic rabbit with a large pendulous abdomen is held in the vertical tail-down position, stagnation of blood in the splanchnic vessels occurs to such an extent that in from fifteen to twenty minutes the animal dies from cerebral anemia. If an abdominal binder is first of all applied, the vertical position will not have the same consequences. This experiment illustrates clearly the possible evil effects that gravity may produce in animals in which no mechanism exists to compensate for it.

Placing an animal such as a dog under light ether anesthesia in the vertical tail-down position produces an immediate fall in arterial blood pressure, as shown in the tracing (Fig. 76), followed by a certain degree of compensation even while the animal is still in the erect position. The extent to which this compensation occurs varies with the depth of the anesthesia. If the experiment is repeated after administering a large dose of chloroform, not only will the initial fall be much greater, but

subsequent compensation will be practically absent. The application of these facts in the operating room will be self-evident.

Leonard Hill has shown that three factors are involved in the compensating mechanism: (1) the tonicity of the abdominal musculature;

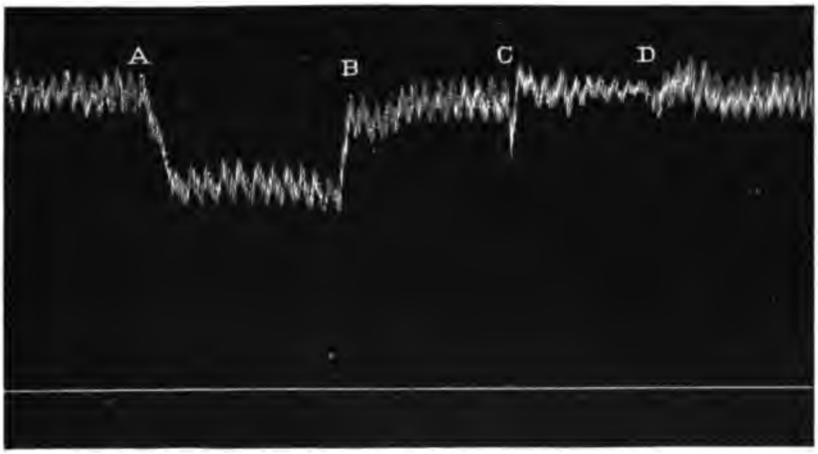


Fig. 76.—Aortic blood pressure, showing the effect of posture: *A*, vertical, head-up; *B*, horizontal; *C*, vertical, head-down; *D*, horizontal. (L.H.)

(2) the tone of the splanchnic blood vessels; (3) the pumping action of the respiratory movements. The importance of the first-mentioned factor can be readily shown by making a crucial incision of the abdominal walls in an animal in the erect position (Fig. 77), and that of

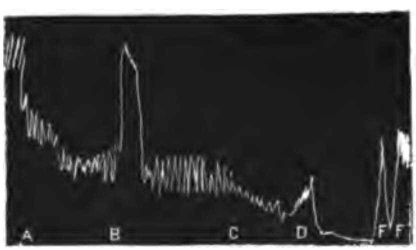


Fig. 77.—Tracing to show the effect of gravity on the arterial blood pressure. At *A*, the animal was placed in the vertical position; at *B*, the abdomen was compressed; at *C*, a crucial incision was made in the abdomen; at *D*, the pleural cavity was opened; at *F*, the animal was returned to the horizontal position. (From Leonard Hill.)

the second factor by cutting the great splanchnic nerves, or the spinal cord. After such an operation, even while in the horizontal position, as we have seen, the blood pressure falls to a considerable extent. If the animal is now placed in the vertical tail-down position, however, it falls

to the zero line and the animal soon dies (Fig. 78). The influence of the third factor is not so great as of the other two, but can be shown by the increased respiratory activity which is likely to develop in the vertical

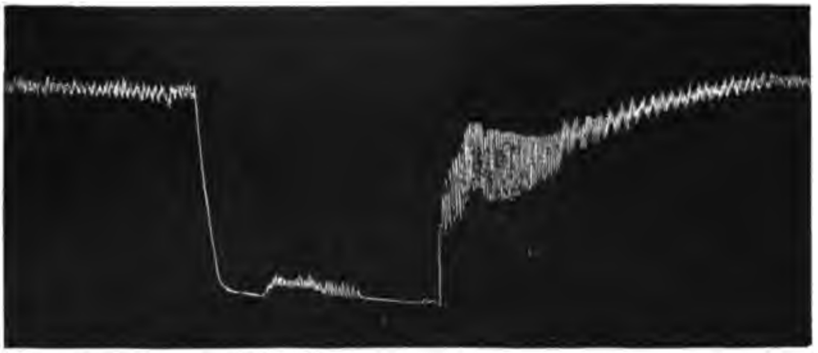


Fig. 78.—The effect of gravity on the aortic pressure after division of the spinal cord in the upper dorsal region. By placing the animal in the vertical feet-down posture, the pressure fell almost to zero, but on returning it to the horizontal posture, the circulation was restored. (From Leonard Hill.)

tail-down position, the anemic condition of the respiratory center being no doubt the cause of the increased respiration.

CHAPTER XXVIII

PECULIARITIES OF BLOOD SUPPLY IN CERTAIN VISCERA

Up to the present we have been considering the circulation of the blood from a general point of view. There are certain organs and tissues, however, in which the general mechanism is altered in order to meet peculiar requirements of blood supply. Thus, it is evident that the brain, incased as it is in the rigid cranium, will be unable to contract and expand as a result of vasoconstriction or vasodilation. On the other hand, we know that the blood supply to this organ does vary considerably from time to time. What is the nature of the mechanism by which such changes are brought about? In the case of the liver the circulation is peculiar on account of the fact that blood is carried to the organ by two vessels, in one of which it is supplied under high pressure and in the other, under low pressure. We must investigate the relationship of these two sources of blood supply. The circulation through the coronary and pulmonary vessels must likewise receive special attention on account of the highly specialized functions of these organs.

THE CIRCULATION IN THE BRAIN

Anatomic Peculiarities

Serious curtailment of the *blood supply* to the brain is guarded against by the existence of the circle of Willis. Besides the four main arteries—the vertebrals and the two carotids—the spinal arteries contribute to the blood supply of the circle, and consequently in certain animals, such as the dog, the four main arteries may be ligated without causing death. In man, however, ligation of both carotids is usually fatal. The free anastomosis displayed in the circle of Willis is not maintained in the case of the arteries which run from it to supply the brain structure. On the contrary, these vessels are more or less terminal in character; that is to say, the capillary system produced by the different vessels does not freely anastomose, so that the obstruction of one vessel, or an important branch, is followed by death of the supplied area. The vessels which go to the pia mater, however, break up into numerous smaller branches, which freely anastomose before entering the brain tissue.

The venous blood is collected by the small, very thin-walled and valveless cerebral veins. These run together to form larger veins discharging into the sinuses, the openings into which are kept patent by the arrangement of dura mater around the orifices. The sinuses exist between the dura and skull and are so constructed that they can not be compressed, particularly those at the base of the brain. From them the blood is conveyed mainly to the internal jugular vein, some of it however escaping by the anastomoses existing between the cavernous sinus and the opththalmic veins, and by the venous plexus of the spinal cord. The most striking peculiarities of the veins are their patulous condition and the absence of valves, so that any change in the blood pressure in the internal jugular vein must be immediately reflected in that of the venous sinuses. This explains why compression of the abdomen

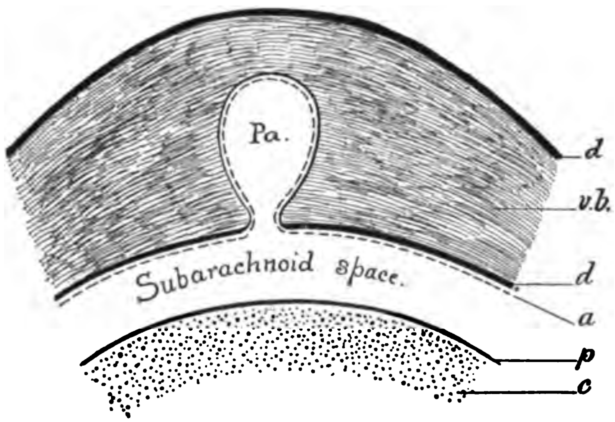


Fig. 79.—Schema to show the relations of the Pacchionian bodies to the sinuses: *d, d*, Folds of the dura mater, inclosing a sinus between them; *v.b.*, the blood in the sinus; *a*, the arachnoidal membrane; *p*, the pia mater; *Pa.*, the Pacchionian body as a projection of the arachnoid into the blood sinus. (From Howell's *Physiology*.)

causes venous blood to flow from an opening made in the longitudinal sinus.

In considering the cerebral circulation, another factor that must be borne in mind is the presence of *cerebrospinal fluid.* This is contained in the subarachnoid spaces of the brain and spinal cord, these spaces, in the case of the brain, being often considerably enlarged to form the cisternæ. The cerebrospinal fluid is also present in the ventricles of the brain, which it will be remembered communicate with the subarachnoid spaces through the foramen of Magendie, etc. It is unlikely that the cerebrospinal fluid is of much importance in connection with the control of the blood supply to the brain tissue. It may be merely a lubricating fluid; at least it is so small in amount (60 to 80 c.c. in man) as to be apparently of little value in bringing about an alteration in brain volume.

Although normally so scanty, its secretion can become remarkably stimulated under certain conditions as in fractures of the base of the skull. Under these conditions in man, it may drain away at the rate of about 200 c.c. a day or more.

The fluid is apparently secreted from the choroid plexus, for when the pathways by which the ventricles communicate with the subarachnoid space are obstructed, it collects in the ventricles, producing internal hydrocephalus. Under certain conditions its absorption is also very rapid, as shown experimentally by the rapidity with which physiologic saline is absorbed when it is injected into the subarachnoid space. This absorption is believed to occur through the Pacchionian bodies, which are minute sac-like protrusions of the arachnoid into the interior of a venous sinus. The membrane that separates blood and cerebrospinal fluid is extremely thin at these places (Fig. 79).

Physical Conditions of Circulation

On account of these anatomic peculiarities, the physical factors controlling the circulation of blood to the brain are considerably different from those obtaining in any other part of the body, with the possible exception of the bones. In other vascular areas, we have seen that, when dilatation or constriction of the vessels occurs, a marked increase or diminution of the volume of the part becomes evident. Such a change in volume is evidently impossible in the case of the brain because of the rigid cranium in which it is contained. In fact, from a physical point of view we must consider the blood vessels of the brain as projecting into a rigid case filled with incompressible material. Under these conditions it is obvious that the vessels as a whole could neither contract nor dilate without some increase or decrease in the volume of the contents of the cranial cavity (Leonard Hill[19]).

Some have thought that the cerebrospinal fluid as it flows into or out of the spinal cord might accomplish this alteration in the cranial contents, but the relatively small amount of available cerebrospinal fluid, the smallness of the openings between the brain and the spinal cord, and the lack of experimental evidence that such changes in volume of cerebrospinal fluid in the spinal cord do actually occur, all stand in contradiction to such a view. However, although the vessels as a whole might not contract or expand, yet some vessels, like the arteries, might contract simultaneously with a corresponding dilatation of other vessels, such as the smaller cerebral veins. In admitting the possibility of some reciprocal relationship between arteries and veins, we must remember that it is only before the well-protected sinuses are reached that a change in the caliber of the veins would be possible. But it is difficult

to see how such reciprocal dilatation and constriction could be of any advantage except perhaps in causing certain areas to receive more blood than others. A reciprocal relationship might also exist between adjacent arterioles as well as between arterioles and veins; when, for example, the arm center becomes active, it is conceivable that its arterioles might dilate at the same moment that those of a neighboring, less active center become constricted. Alterations obviously might occur without causing any perceptible change either in the volume of the brain as a whole or in the condition of venous flow.

In consideration of these factors, most observers are agreed that the total volume of blood in the brain must be constant at all times (Monro and Kellie doctrine). Alteration of blood supply can, however, still be brought about by changes in the velocity with which the blood traverses the vessels. When more blood is required in the brain to supply the increased metabolism which we must presume accompanies heightened mental activity, it is not accomplished as in other parts of the body by an increase in the capacity of the vessels as compared with those of other vascular areas, but by a hurrying up of the circulation through vessels whose caliber remains unaltered.

The main factors determining the velocity of bloodflow through the brain must, therefore, be dependent upon changes occurring elsewhere in the vascular system, a conclusion for which there is abundant experimental evidence. Of the many ingenious methods that have been devised to secure this evidence, we will cite but one in this place. Records are taken of changes in: (1) the *venous blood pressure* of the brain by connecting a cannula either with the vein immediately after leaving the skull or, better still, with the *torcular Herophili;* (2) the *brain volume,* by connecting a very sensitive receiving tambour with a trephine hole in the cranium so that its open end lies against the pia mater.* Although, as we have seen, while incased in the rigid cranium the brain volume can not change to any degree, yet this will occur when a portion of the cranium is removed, so that pulsations corresponding to those in the blood vessels will be observed; (3) the circulatory conditions elsewhere in the body, by taking arterial and venous pressures and plethysmograms. The results in a normal animal show the following points (see Fig. 80): (1) The tracings of the arterial blood pressure (*A*), the brain volume (*C*) and the intracranial venous pressure (*C*) have exactly the same contour—that is, the respiratory and the cardiac waves in all three of them are identical. The venous blood as it flows into the jugular veins also pulsates in

*This receiving tambour really consists of a brass tube of the same diameter as the trephine hole, into which it is tightly fitted. The brass tube is closed at its inner end by thin rubber membrane, and its outer end is connected with the receiving tambour.

unison with the artery. (2) Any change in the blood pressure of the systemic venous system is immediately reflected in the blood pressure of the sinuses of the brain and in the brain volume (not well shown in accompanying tracing). (3) A change never occurs in the vessels of the brain which can not be accounted for by some change occurring

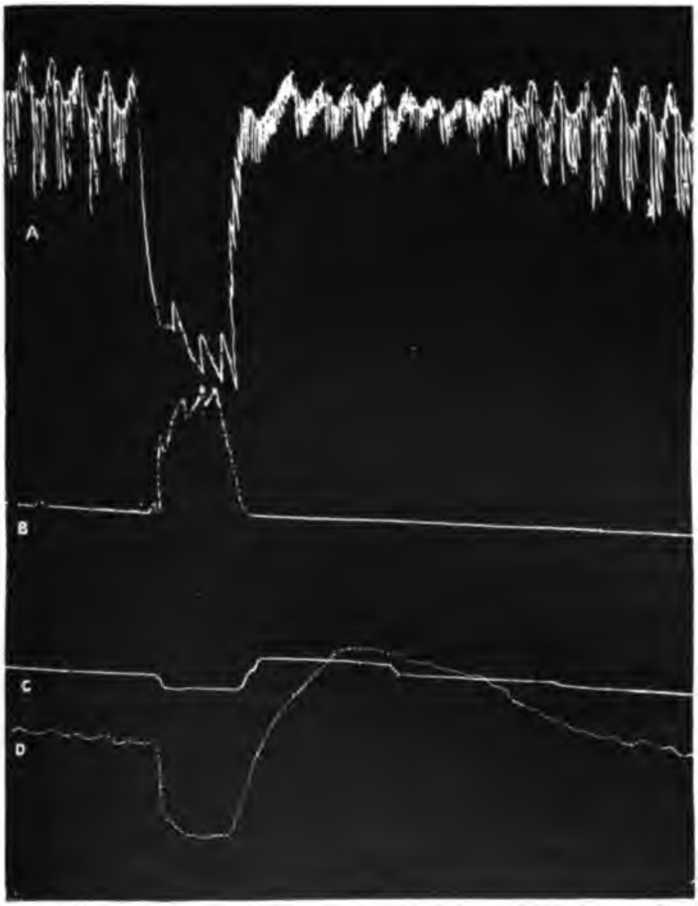


Fig. 80.—To show simultaneous records of the arterial blood pressure (*A*), the venous pressure (*B*), the intracranial pressure (*C*), the pressure in the venous sinuses (*D*). The fall in arterial pressure produced by stimulation of the cerebral end of the vagus will be found to cause a fall of intracranial and cerebral venous pressure, accompanying that in the arteries, but a rise in that of the venous system. (From Leonard Hill.)

elsewhere in the vascular system outside the cranial cavity. This result is important because it shows that there can not be vasomotor nerve control of the brain vessels.

Taking into consideration not only the results of such experiments, but also the peculiar physical conditions existing in the cranial cavity,

we must conclude that changes in blood supply depend on changes in the velocity of the bloodflow, and that such alterations in velocity are dependent upon changes occurring in the aortic and more especially in the vena-cava pressure. When the aortic pressure rises, more blood will flow into the cerebral arteries and move along them at an increased velocity, the increased pressure probably causing a moderate degree of passive dilatation, to allow extra room for which the numerous *small* cerebral veins become compressed. This compression of the veins probably does not obstruct the greater flow of blood through them, because, taken as a whole, they are ordinarily much more capacious than need be. On the other hand, if the aortic pressure should remain constant, but that in the vena cava increase, then there would be obstruction to the passage of blood in the intracranial arteries, and consequently a diminished velocity of flow.

Vasomotor Nerves

It might be inferred that, since the bloodflow through the cerebral vessels is mainly dependent on vascular conditions elsewhere in the body, there would be no need, as in the vessels of other vascular areas, for vasomotor fibers. Histologists have, however, discovered the presence of such fibers, and it has become necessary for the physiologist to find out if they are really of importance in connection with the regulation of the blood supply to the brain. Even if it is admitted that the arterioles could not contract or expand as a whole without producing local changes in venous pressure or cranial volume, it is yet of course always possible, as has already been pointed out, that one set of arterioles might contract at the same moment that another set expanded.

That the vessels can undergo a process of constriction has been shown by experiments in which the volume of outflow from the vessels of the brain was measured in perfused preparations of brain. When epinephrine was added to the perfusion fluid, curtailment of outflow was observed to occur (Wiggers). Since this drug causes constriction of vessels only when these are supplied with constrictor fibers (see page 736), the conclusion may be drawn that the cerebral blood vessels do contain such nerve fibers. Nevertheless, the local vasomotor control of the cerebral blood vessels can not have the significance in connection with changes in blood supply that it has for other vascular areas (Hill and Macleod[20]). No doubt nerve fibers are present in the cerebral blood vessels, and presumably under certain conditions they are capable of causing the blood vessels to undergo alterations in caliber, but it is impossible *to see of what real value this can be under normal conditions.*

Intracranial Pressure

One word more with regard to what is known as *intracranial pressure,* that is, the pressure in the space between the skull and the brain. Under ordinary conditions it must be equal to that in the cerebral capillaries, and may be measured by connecting a sensitive manometer with a tube screwed into the cranium as described above. It has been found to vary from 0 mm. Hg in a man standing erect to 50-60 mm. Hg in a dog poisoned by strychnine. It becomes increased, not only by compression of the veins of the neck and by an increase in general arterial pressure, but also in pathologic conditions, such as hydrocephalus. A new growth in the brain, if it occupies more space than the tissue which is destroyed, exerts pressure on all parts of that region of the cranial cavity, but this pressure may not be transmitted equally throughout the cranial contents, for the falciform ligaments and the tentorium support a part of it, thus directing the spread of pressure along certain pathways. The structures at the base of the brain, the optic nerves, the veins of Galen and the Sylvian aqueduct are most affected in this way. If the pressure is rapidly applied, however, it may rise throughout the cranial contents. In such cases the pressure is, of course, circulatory in origin, since immediately after death from cerebral tumor the intracranial pressure is not found to be raised.

The major *symptoms of cerebral compression* are no doubt due to anemia of the medulla oblongata, which may be the result either of pressure applied locally in the bulbar region, where the presence of a very small foreign body or only trivial tumor formation is sufficient to destroy life, or of pressure transmitted from the cerebral cavity, in which case, on account of the support offered by the tentorium, a much larger growth is required to affect the medulla. Internal hydrocephalus produced by blocking of the aqueduct of Sylvius and the veins of Galen causes the greatest rise in intracranial tension, and may affect the medulla, because the brain is driven downwards so as to pinch the bulb against the occipital bone. It must be emphasized that it is not the pressure *per se* that causes the symptoms, but the attendant anemia, the symptoms of acute cerebral anemia and of compression being identical (Leonard Hill[19]). To relieve the compression, trephining is the common practice. The trephine hole should be as large and as near to the source of compression (tumor, etc.) as possible.

CIRCULATION THROUGH THE LUNGS

The pulmonary or lesser circulation, as it is called, is quite different from the systemic circulation. In the first place, because the pressure

in the pulmonary arteries does not amount to more than about 20 mm. Hg, or about one-sixth of that of the systemic arteries, the peripheral resistance in the blood vessels of the lungs is much less than that of the body in general. This lower resistance is owing partly to the large diameter of the arterioles and the small amount of muscular fibers in their walls, and partly to the fact that the capillaries are held constantly in a somewhat dilated condition on account of the subatmospheric pressure in the thorax (see page 306).

Another peculiarity of the pulmonary circulation is that the caliber of the vessels is to a very large extent dependent upon the changes that occur in the intrathoracic pressure with each inspiration and expiration. They become dilated on inspiration and contracted on expiration. The extent to which these respiratory changes affect the amount of blood contained in the lungs, is very considerable. At the height of inspiration it is computed that a little more than eight per cent of the whole blood in the body is contained in the lungs, whereas on expiration it diminishes to between five and seven per cent.

A third peculiarity is that the pulmonic blood vessels are not supplied with vasomotor nerve fibers—at least with such as can readily be demonstrated. It is said that, when the pulmonary vessels are perfused and the outflow measured, a diminution in the latter is found to occur when epinephrine is added to the injection fluid—a result which is, however, denied by certain investigators. Changes in the bloodflow have not been observed to occur when the vagus or sympathetic nerve fibers running to the lungs are stimulated. In short, the conclusion which we must draw is much the same as that for the blood vessels of the brain—namely, that although, as a result of the epinephrine experiment, we must admit that a vasomotor supply *may* possibly be present, yet it is one which can be of no significance under normal conditions.

When there is obstruction to the outflow of blood from the left ventricle, as, for example, in cases of high aortic pressure, the blood is not entirely discharged with each beat of the left ventricle, and therefore dams back through the left auricle into the lungs. On account of the marked distensibility of the pulmonary capillaries, a large amount of this blood may collect there and thus make the lungs serve as a kind of reservoir of the heart. When the capacity of this reservoir has, however, been overstepped, an increased peripheral resistance will come to be offered to the movement of blood in the pulmonary arteries, the pressure in which will consequently rise and sooner or later interfere with the discharge from the right ventricle, causing as a result a stagnation of blood in the systemic veins, and a consequent increase in vol-

ume of such viscera as the liver and kidneys. The same changes will obviously also supervene when there is regurgitation of blood from the left ventricle to the left auricle, as in cases of mitral insufficiency.

CIRCULATION THROUGH THE LIVER

The liver is the only gland in the body receiving both venous and arterial blood, the former being supplied to it at a very low pressure by way of the capacious portal vein, and the latter at very high pressure by the strikingly narrow hepatic artery. Except for the relatively small amount of blood which is supplied to the walls of the blood vessels and the biliary ducts, none of the hepatic artery blood mixes with that of the portal vein until the vessels enter the hepatic lobules. Beyond this point the two blood streams mix and the combined stream is drained away by the sublobular and hepatic veins.

Methods of Investigation

To study the relative importance of these two sources of blood supply, and also to investigate the manner in which the latter is controlled, the most satisfactory method has consisted in measurements of changes in volume flow rather than in those of changes in pressure. The volume-flow measurement has been made either by connecting stromuhrs (page 207) to the hepatic artery or portal vein, or by measuring the outflow of blood from the hepatic vein into the vena cava, first with both inflow vessels intact, and then with one of them ligated. An objection to the first (the stromuhr) method is the possible interference with bloodflow or blood pressure produced by inserting the stromuhr into the entering vessels, and also the fact that simultaneous measurement of the flow in both vessels can not be made satisfactorily.

To measure the outflow from the hepatic veins, the aorta is ligated below the celiac axis and a wide cannula is inserted into the central end of the vena cava below the level of the liver, a loose thread being placed around this vessel just above the diaphragm. By pulling on this thread the vena cava becomes obliterated, and the blood from the hepatic veins is therefore diverted into the cannula, through which it flows into one end of a vessel shaped somewhat like a sputum cup (the receiver), the other end being connected by tubing with a piston recorder, from the movement of which the volume of blood flowing into the receiver can readily be computed. To measure the flow of blood, a clip on the tube of the receiver is removed at the same moment that the thread around the vena cava above the diaphragm is tightened, and when the receiver has filled with blood, this thread is again loosened

and the receiver tilted up so that the blood flows at low pressure back into the circulation. The receiver being of known capacity, the length of time it takes the blood to fill it as determined by the piston recorder, furnishes us with the necessary data from which to calculate the rate of flow. The receiver is chosen of such a size that it takes only a few seconds to fill, the diversion of blood into it not causing any material fall in arterial pressure. The observations are repeated frequently.

Results.—By the use of these methods it has been found that the total mass movement of blood to the liver of the dog varies between 1.46 and 2.40 c.c. per second for 100 grams of liver. Considerable changes may occur in the arterial pressure without affecting the liver flow. When the hepatic artery is occluded, the flow diminishes by about 30 per cent, or conversely, when the portal vein is obstructed but the hepatic artery left intact, by about 60 per cent, indicating that about one-third of the total bloodflow through the liver is contributed by the hepatic artery and two-thirds by the portal vein. Some blood, however, gains the liver through anastomotic channels between it and the diaphragmatic veins.

The relative supply by the two vessels is subject to various conditions. That through the hepatic artery, for example, may be very considerably altered on account of vasoconstriction in this vessel, for its walls can easily be shown to be liberally supplied with vasoconstrictor fibers carried by the hepatic plexus. This can be demonstrated by the rise in blood pressure which occurs in a branch of the hepatic artery during stimulation of the plexus. On the other hand, alterations in the bloodflow in the portal vein can not be brought about by active constriction or dilatation of the intrahepatic branches of this vessel, no active vasomotor fibers having been demonstrated by stimulation of the hepatic nerves, although, as in the case of the brain and lung blood vessels, a certain amount of constriction may occur under the influence of epinephrine.

The bloodflow through the portal vein is dependent on changes occurring at either end of the distribution of the vessel, that is, changes occurring in the liver itself or in the intestine. Of these factors the latter is no doubt the more important, an increase not only in portal blood pressure but also in portal bloodflow being readily produced by dilatation of the splanchnic blood vessels; for example, as the result of section of the splanchnic nerve. Alterations in portal bloodflow brought about by changes in the caliber of the vessels in the liver itself are partly dependent upon changes in the branches of the hepatic artery. Let us consider briefly how this may be brought about. At the point where the portal and hepatic arteries come together—that is, at the in-

trahepatic capillaries—the pressure of the blood in them must become equal, which means that in its course through the interlobular connective tissue, the branches of the hepatic artery must offer much resistance to the blood flowing through them. This frictional resistance resides in the hepatic arterioles, and since these are richly supplied with constrictor nerves, great variation in hepatic inflow becomes possible. These changes will affect the degree of tension of the interlobular connective tissue in which the arterioles lie. In this tissue, however, also lie the thin-walled branches of the portal vein. When therefore the tension of this tissue becomes greater, as a result, for example, of vasodilatation in the hepatic artery, the portal vein radicles will become compressed and the bloodflow along them impeded. Conversely, when vasoconstriction occurs in the hepatic arteries, the congestion of the connective tissue becomes diminished, the veins dilate, and the blood flows through them more readily (Macleod and R. G. Pearce[21]). Experimental evidence in support of the above view is furnished by observing the outflow of blood from the liver before and during stimulation of the hepatic plexus. The first effect is an increase in the outflow, which very soon returns to its original amount, even though the stimulation of the plexus is kept up during the experiment. This return to the normal flow must indicate either that the constriction of the hepatic artery has not been maintained, or that it has been maintained but is accompanied by a compensatory increase in the flow through the portal vein. As a matter of fact, we know that the hepatic artery remains constricted as long as the hepatic plexus is stimulated, indicating that the congestion of the connective tissue in which the venules lie has become reduced to such an extent, as a result of the constriction, that these open up and permit the blood to flow through them more readily. The initial increase in outflow immediately following upon stimulation of the hepatic plexus, is no doubt caused by the squeezing out of the blood already in the hepatic vessels, and it is a result which is often observed in other organs during stimulation of vasoconstrictor nerve fibers.

THE CORONARY CIRCULATION

We have already studied the effect produced on the heartbeat by interfering with the flow of blood in the coronary vessels, and it remains for us to study: (1) peculiarities in the bloodflow through them, and (2) whether this bloodflow can be altered by dilatation or constriction of the vessels brought about through nerves. With regard to the *peculiarities of bloodflow,* it may be stated that there is said to be two periods in each cardiac cycle during which an increase takes place in the mass

movement of blood in the coronary vessels—namely, at the beginning of systole, and again at the beginning of diastole. Nevertheless the pressure pulse has the same contour in the coronary as in the systemic circulation. (W. T. Porter.[22]) During systole the intramural branches of the coronary artery are compressed and the blood pressed out of them. This emptying of the vessels favors the flow of blood through the heart walls.

Regarding the presence of *coronary vasomotor nerves,* there is at present a certain amount of doubt. When strips of the coronary artery are suspended in a solution of epinephrine, they undergo relaxation instead of contraction. On the assumption that the action of epinephrine on blood vessels is the same as that of stimulation of the vasoconstrictor fibers, this result has been taken as evidence of the absence of such fibers and the possible presence of vasodilator fibers. A somewhat similar type of experiment has been performed by injecting epinephrine into the fluid used to perfuse the excised mammalian heart, with the result that, when such injections are made into a heart that is not beating, evidence of vasoconstriction is obtained, whereas when injected into a beating heart, dilatation occurs. This latter result may, however, be owing to the action of the epinephrine in stimulating the cardiac contractions. Other observers, however, deny that the injection of epinephrine into the coronary circulation has any influence upon the outflow of the perfusion fluid. Taking the result of these observations as a whole, we may at least conclude that epinephrine does not produce the same marked vasoconstriction that it produces in other blood vessels—a fact, which, as already stated, may be taken advantage of in bringing about the rise in coronary pressure that is necessary for successful resuscitation of the heart.

Attempts to demonstrate the presence of vasomotor fibers by electrical stimulation of the vagus or sympathetic nerve have yielded results which are quite inconclusive, although some observers assert that the vagus nerve carries vasoconstrictor fibers to the coronary vessels, and that the sympathetic carries vasodilator.

CHAPTER XXIX

CLINICAL APPLICATIONS OF CERTAIN PHYSIOLOGIC METHODS*

In the following chapters a brief account will be offered of the clinical use of the electrocardiogram, of polysphygmograms, and of bloodflow measurements. This is done to show how physiologic technic is being employed for the accurate investigation of cardiovascular disease.

ELECTROCARDIOGRAMS

To observe the electrical change produced by the spread of the excitation wave over the heart from auricles to ventricles, it is not necessary to place the electrodes directly on the heart, but, as already hinted, we may follow the electrical change by leading off from electrodes applied to the surface of the body. From such electrocardiographic tracings extremely important facts concerning the propagation of the heartbeat may be ascertained. In order to make an observation the hands and the left foot are each placed in a solution of sodium chloride contained in porous jars, immersed in larger vessels containing a saturated solution of $ZnSO_4$ and zinc terminals.† An arrangement like that in Fig. 81 may also be used. By manipulation of suitable keys the extremities may then be connected with the electrocardiograph in the following manner: Lead 1, right arm and left arm; lead 2, right arm and left leg; lead 3, left arm and left leg. Through lead 1, the current acting on the galvanometer will be that produced more especially at the base of the heart. Through lead 2, the current will pass through the long axis of the heart, and through lead 3, it will pass mainly along its left border.

When any pair of leads is connected with the galvanometer, it is observed that the string is deflected to one side owing to electrical currents arising from the skin. Before taking a record of the cardiac movements of the string, it is necessary to *compensate* for this skin current by introducing into the circuit in the opposite direction the re-

*A certain amount of repetition of matter previously discussed has been found advisable in these chapters for which the indulgence of the reader is requested.

†It is really unnecessary to use the so-called nonpolarizable electrodes. Glass vessels containing 20 per cent NaCl solution with the zinc plates dipping into them are quite satisfactory.

quired amount of current, called the compensating current, to bring the strong shadow back to the zero or midposition. In order that the record obtained may be quantitative in character, it is further necessary that the movement of the string be *standardized.* This is done by ascertaining to what extent the string moves when a current of known voltage is sent through it and by altering the tension of the string so that one millivolt of current causes an excursion of one centimeter of the string shadow on the photographic plate. It would take us beyond the

Fig. 81.—Electrocardiographic apparatus as made by the Cambridge Scientific Materials Co. Contact electrodes are shown, but the immersion electrodes described in the context are preferable.

confines of this volume to go in any greater detail into the technic involved in taking electrocardiograms, but it may be said that this is by no means difficult, provided the instructions which are supplied with the instrument are carefully followed. In practice the taking of electrocardiograms is indeed quite a simple matter, and the extremely important information which they give us concerning the mechanism of the heartbeat and the evidence of myocardial disease should make their employment a universal practice in all cardiac clinics. Some of these clinical applications are described elsewhere (page 266).

What particularly interests us here is the *contour of the electrocardiogram in a normal person* (Fig. 82). It will be observed that there are three waves above the line of zero potential and two waves below it. They have been lettered from before backward, P, Q, R, S, and T, and in all such records when correctly obtained, the waves above the line of zero potential indicate that the base of the heart is negative to the apex. The exact cause of each wave has been ascertained by taking simultaneously with the electrocardiogram a record of the mechanical changes occurring in the heart during each cardiac cycle. Such records

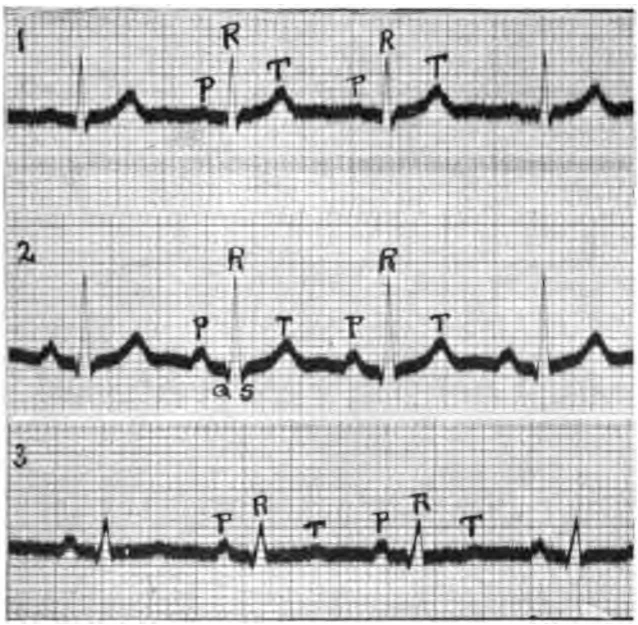


Fig. 82.—Normal electrocardiogram. Leads 1, 2, 3. Note that the height of the *R* deflection in lead 3 equals the difference between the height of R_1 and R_2.

have been secured by taking intracardiac pressure curves with the results as shown in Fig. 83. The top curve represents auricular and the second one ventricular pressure, whereas the lowest is an electrocardiogram. It will be observed: (1) that the P-wave occurs just antecedent to contraction of the auricles; (2) that the small positive wave, Q, which is absent in these tracings, must occur just before the beginning of the contraction of the ventricles; (3) that the negative wave, R, occurs just before and during the early part of ventricular systole—that is, during the presphygmic period; and (4) that the long upward wave, T, culminates at the moment the ventricle begins relaxing.

Although such comparisons give us considerable insight into the cause of several of the waves, there yet remain certain peculiarities of the electrocardiogram to be considered. These are: (1) the cause of the slight positive wave, Q; (2) the cause of the positive wave, S; (3) the cause for the period of equal potential at the base and apex during ventricular systole indicated by the portion of the curve between S and T; (4) the cause for the negative wave, T. To solve these problems it is necessary to compare electrocardiograms taken from the surface of the body with those from electrodes placed directly on the base or apex of the ventricle of the exposed heart.

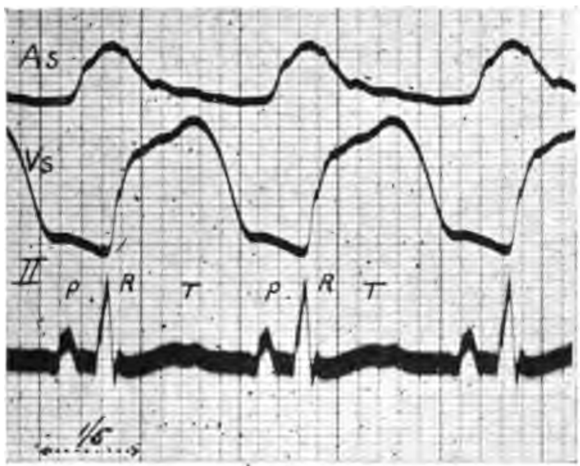


Fig. 83.—Electrocardiogram (dog) taken simultaneously with curves from auricle and ventricle. It will be observed that wave *P* slightly precedes auricular systole and that wave *R* occurs just before the presphygmic period starts in the ventricle. (From Lewis.)

The Ventricular Complex

In view of the nature of the electric change which occurs in a strip of denervated muscle when a wave of contraction passes along it (page 188), the simplest interpretation of the ventricular part of the above curve is that the contraction must pass into the ventricle at a little distance from the base, thus causing the latter, for a moment of time, to be positive to the rest of the ventricle, and accounting for the slight downward wave, Q. Immediately after this the base of the ventricle becomes negative to the apex, giving us the marked upward wave, R, which however lasts for but a short period of time, being followed by an interval during which the base and apex are of the same electrical potential (horizontal part of wave between R and T). Finally the base again becomes negative to the apex, thus accounting for the smaller upward

wave, T. The cause of the occasionally observed downward wave, S, following R, is obscure.

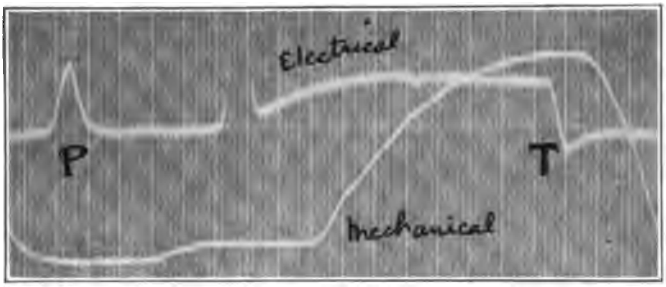
The most significant fact in the electrocardiogram is therefore that *the base is negative to the apex at the beginning* (*R-wave*) *and again at the end* (*T-wave*) *of the ventricular contraction.* How may this be explained? When electrocardiograms are taken through electrodes placed directly on the base and apex of the ventricle of the exposed heart, it has been found that the contour of the electrocardiogram is like that which is obtained from a strip of muscle when a wave of contraction passes along it: it is diphasic in character (page 188), a result which may be interpreted as indicating that the wave of contraction starts at the base and ends at the apex. This rules out the explanation, at one time suggested for the T-wave, that the wave starts at the base, then proceeds to the apex, and finally ends at the base, following the disposition of the muscular fibers of the ventricle in a folded or loop form, with the bend of the loop at the apex and the free ends at the base. Although the explanation seemed at first to conform with the embryologic fact that the heart is developed from a folded tube, it can not hold, as has been shown by observing the course of the excitation wave secured through electrodes placed at various points on the surface of the exposed ventricle (page 194).

The explanation which is accepted by the majority of observers at the present time is to the effect that *the T-wave is caused by the longer continuance of the electric change at the base of the ventricle than at the apex.* To test this hypothesis the crucial experiment would evidently be to see whether a T-wave could be induced in an electrocardiogram, such as that of the frog ventricle, in which no T-wave exists, by hurrying up the contraction process at the apex without affecting it at the base. This can be done by local warming of the apex, or by applying the ventricular electrode at varying parts of the ventricle in an excised heart beating in Ringer's solution of relatively high H-ion concentration. Mines showed that under these conditions a typical T-wave appears in the electrocardiogram, as shown in Fig. 84.*

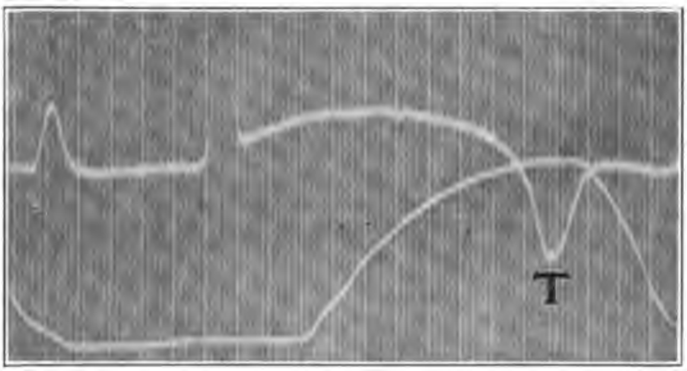
The existence of the small Q-wave, indicating that the contraction does not really start from the base, conforms with the observation that the Purkinje system of fibers ends about the papillary muscles, which therefore would be the first to contract, and with the observations of Lewis, already alluded to above, on the appearance of the negative variation on the surface of the exposed heart.

The most important clinical application of the electrocardiogram is

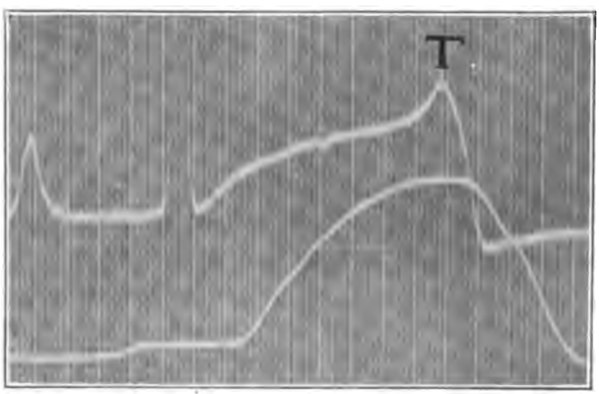
*This tracing was found among those left by Professor Mines of McGill University, and for permission to use it the author is indebted to the authorities of that institution.



A.—Normal



B.—Apex cooled



C.—Apex warmed

Fig. 84.—Records of electrocardiogram and movement of ventricle of frog showing that when the apex is warmed a typical T-wave appears in place of a wave in the opposite direction appearing when the apex is cooled. (From Mines.)

undoubtedly in connection with the determination of *the rate of transmission of the excitation wave* from auricle to ventricle; thus, the P-R interval, as it is called, indicates the time taken for the impulse to travel from the sinoauricular to the auriculoventricular node and bundle. In delayed transmission this interval becomes abnormally long. Obviously also conditions of *heart-block,* of *auricular fibrillation,* or of *auricular flutter* will be immediately revealed by the electrocardiogram. The interpretation of abnormalities in the contour of the ventricular portion of the curve is, however, not so easy a matter, and should never be undertaken unless curves from the three leads have been secured, for it will be found that the corresponding electrocardiograms differ from one another in detail; for example, the R-wave is usually most prominent in lead 2, although sometimes it is more prominent in lead 3. T is always upright in normal individuals in curves taken from lead 2, but it is not infrequently inverted in those of lead 3, and may show partial inversion in those from lead 1. The Q-R-S group is often of peculiar contour in curves from lead 3. These variations are possibly dependent upon the relative preponderance of the musculature in the left and right ventricles, for it is evident that the amount of muscle included in the pathway between the two leads will vary.

CHAPTER XXX

CLINICAL APPLICATIONS OF CERTAIN PHYSIOLOGIC METHODS (Cont'd)

CLINICAL APPLICATIONS OF ELECTROCARDIOGRAPHY

The Electrocardiogram in the More Usual Forms of Cardiac Irregularities

BY R. W. SCOTT

The principle of the application of the string galvanometer to the study of cardiac irregularities has been indicated. It is our object here to outline some of the more common forms of irregular heart action, with a brief description of the abnormalities in the electrocardiogram resulting therefrom. For the sake of comparison a normal electrocardiogram is shown in Fig. 82. The cause and relationship of the various deflections have been explained (see page 262).

Sinus Arrhythmia.—This irregularity is seen commonly in children and young adults, and is without pathologic significance. The electrocardiogram presents the normal deflections and shows by the varying spaces between the P deflections that the cardiac impulse has been generated at slightly irregular intervals.

Sinus Bradycardia.—The electrocardiogram in a simple case of sinus bradycardia is usually normal, except that the deflections occur at an unusually slow rate (Fig. 85). This indicates that the cardiac impulse is built up at a slow rate, but when generated it evokes a normal auricular and ventricular contraction.

The Extrasystole.—The extrasystole may be either auricular or ventricular in origin. Occasionally a rare type is seen in which the impulse arises in the junctional tissues between the auricle and ventricle. When the focus of impulse production is at or near the *sinoauricular node,* the resulting electrocardiogram complexes are practically normal. If, however, the seat of impulse formation is removed from the S-A node, the P deflection may be distorted or actually inverted, followed by a normal Q-R-S-T complex (Fig. 86).

In the case of *ventricular extrasystole,* the cardiac impulse originates in either the right or the left ventricle. This abnormal site, together

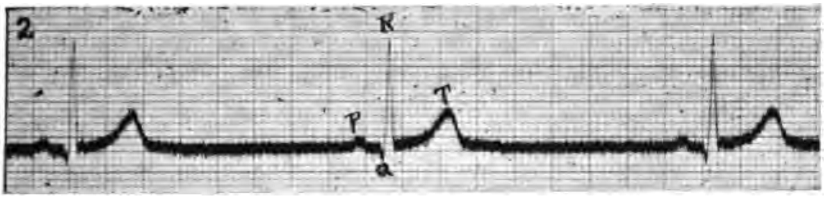


Fig. 85.—Sinus bradycardia. Rate 32 per minute. Note the normal appearance of the electrocardiogram. P-R interval = .17 seconds.

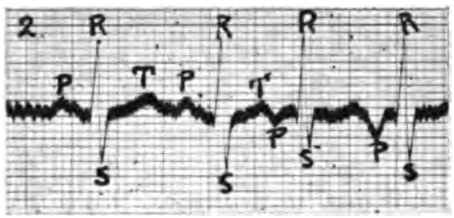


Fig. 86.—Auricular extrasystole. Two auricular extrasystoles following two normal complexes. Note the ectopic origin of the extrasystoles indicated by the inversion of P.

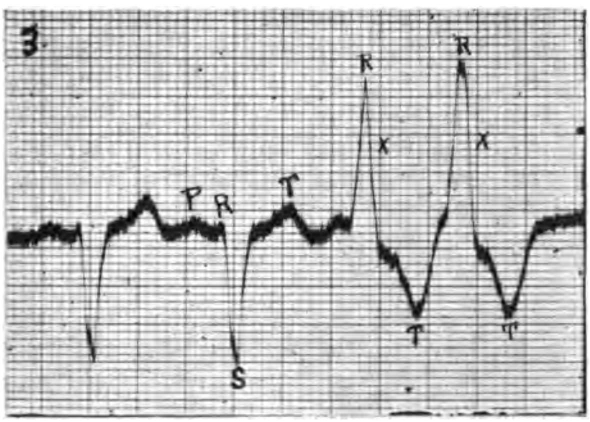


Fig. 87.—Ventricular extrasystoles arising in the right ventricle.

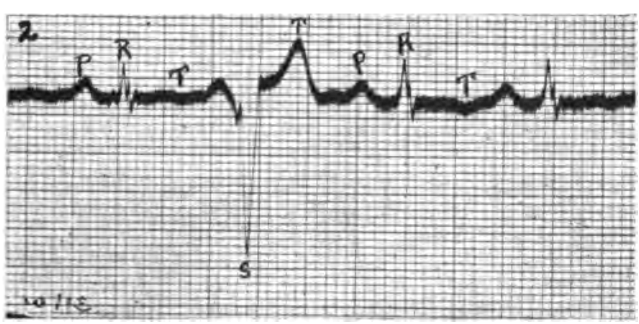


Fig. 88.—Ventricular extrasystole arising in the left ventricle.

with the path which the impulse takes, produces a much greater difference of electric potential than is seen in the normal electrocardiogram. When the impulse arises in the right ventricle near the base, the prin-

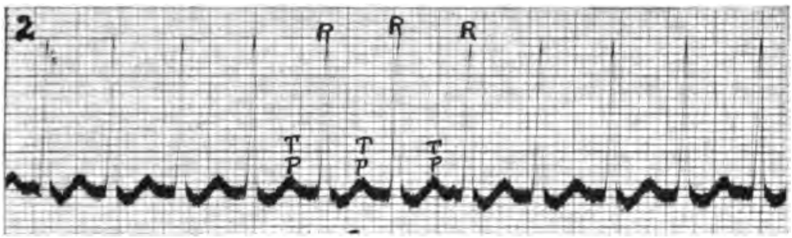


Fig. 89.—Paroxysmal tachycardia. Auricular origin. Note that the *P* deflection falls back on *T*. Rate 200 per minute.

cipal R deflection is upwards in both leads 1 and 2. Arising near the apex, the principal R deflection is up in lead 1 and down in lead 2. Two extrasystoles both arising in the right ventricle are shown in Fig. 87.

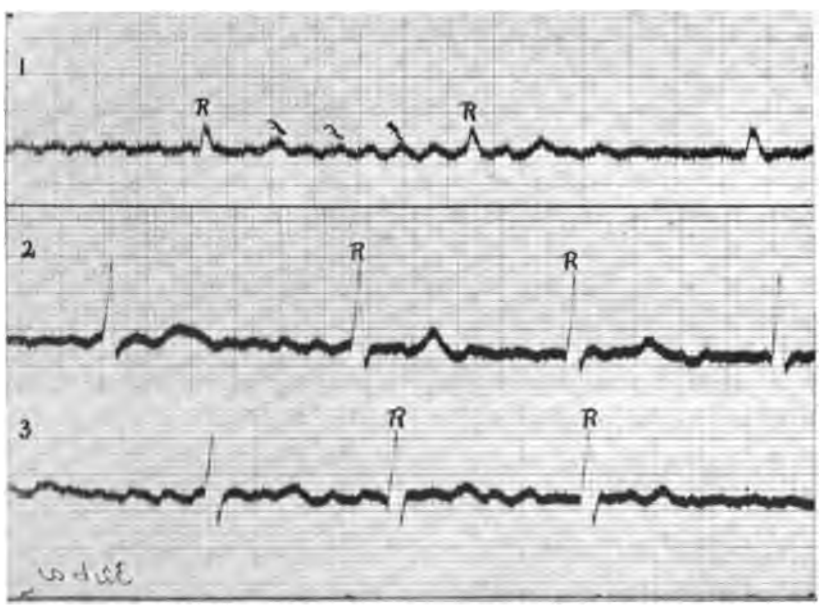


Fig. 90.—Auricular fibrillation. Leads 1, 2, 3. Note the coarse fibrillation waves between the *R* peaks, and the absence of any *B* deflections in relation to *R*. Also the unequal spacing of the *R* deflections.

In the case of the left ventricle, a basal impulse gives a downward principal deflection in lead 1 and up in lead 2. When the aberrant focus is located near the apex of the left ventricle, the principal deflec-

tion is down in both leads 1 and 2. Any one or several of the general types of extrasystole may occur in the same patient. Fig. 88 shows an extrasystole originating from the left ventricle.

Paroxysmal Tachycardia.—Electrocardiographic records taken in the interval between the paroxysms may appear normal. During the tachycardia the records normally show only two deflections, R and a combination of T and the succeeding P (Fig. 89). If the paroxysm is of auricular origin, the P deflection may be inverted, indicating that the new focus of impulse production is located at some other site than the sinoauricular node. Rarely the new focus may be in the ventricles. Records taken during the paroxysm may show a rapid succession of deflections, simulating isolated ventricular extrasystoles.

Auricular Fibrillation.—The electrocardiogram in auricular fibrillation shows three distinctive features:

1. Absence of the P deflections typical of auricular contractions.
2. The ventricular complexes (Q-R-S-T waves) occur in irregular sequence and may vary in height.
3. The presence of small irregular oscillations best seen between the ventricular complexes. A typical tracing of this condition is shown in Fig. 90.

The dependence of the P-wave upon auricular contraction has been indicated (page 261). Its absence in auricular fibrillation is accounted for by the fact that the individual muscle fibers of the auricles contract independently of one another, so that some fibers are in a state of contraction while others are relaxed. This renders impossible a coordinate contraction of the auricle as a whole.

The multiple impulses from the fibrillating auricles reach the ventricles and evoke a contraction provided the ventricle is not already in a state of contraction (refractory period, page 178). These irregular ventricular responses will of course produce unequal spacing of the ventricular complexes in the electrocardiogram. The variations in the height of the R deflections is thought to be due to the distortion caused by the superimposition of the small waves representing auricular activity. These small waves must occur throughout the whole cardiac cycle, but are more or less masked by the ventricular complexes, appearing as separate oscillations only during diastole.

Auricular Flutter.—Auricular flutter was discovered by the electrocardiograph, and it is practically impossible to make a diagnosis of this condition without the use of the string galvanometer. The auricular deflections are usually rhythmic and in the average case vary in rate from 200 to 350 per minute. The initial deflection of P may be base negative or apex negative—up or down—depending on the site of the

origin of the auricular impulse (when arising from some other source than the S-A node the impulse is said to be ectopic). Usually a regular succession of P deflections can be traced throughout the record (Fig. 91).

Since it is impossible for the ventricle to respond to all the impulses coming from the auricles, a condition of partial heart-block obtains (2:1—3:1—4:1, etc.). The ventricular complexes will occur regularly except when a 3:2 rhythm exists.

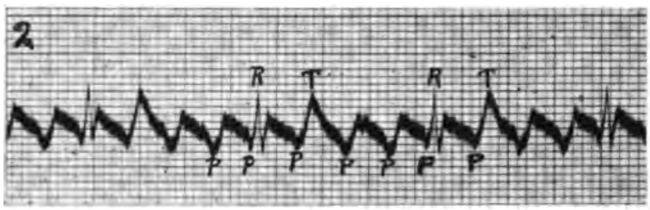


Fig. 91.—Auricular flutter. Auricular rate 300. Ventricular rate 80. Note the inversion of the *P* deflections.

Usually the ventricular complexes are such as to indicate that the stimulus arose in the auricle (supraventricular). The height of the individual deflections Q-R-S-T may vary, depending on the predominance of a right or left ventricular hypertrophy.

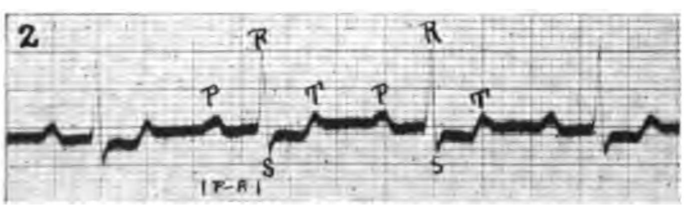


Fig. 92.—Delayed conduction. Note the normal appearance of the electrocardiogram except for the prolongation of the *P-R* interval, which measures .23 seconds.

Heart-block.—There are three degrees of severity in heart-block: (1) delayed conduction, (2) partial dissociation, and (3) complete dissociation.

Any one of these conditions may be present in the same patient at successive intervals.

Delayed Conduction.—When the conducting tissues of the heart are so affected as to cause an abnormal prolongation of the P-R interval, the condition is called delayed conduction. The ventricles respond to each stimulus originating at the sinus node, but the time required for the impulse to pass through the conducting tissues is longer than normal,

In a simple case the electrocardiogram may appear perfectly normal, but when the P-R interval is measured accurately, it will be found to be lengthened beyond the extreme limits of the normal (0.20 seconds) (Fig. 92).

Partial Dissociation.—In the typical case of partial dissociation the

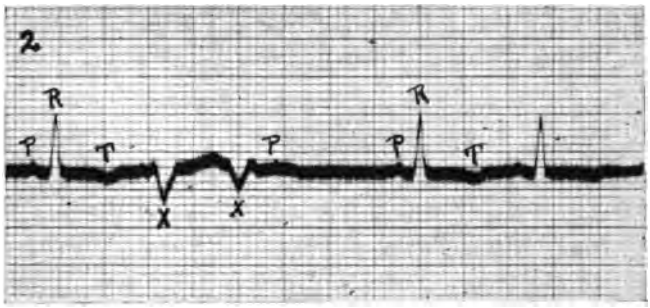


Fig. 93.—Partial dissociation. Note the failure of ventricular response following the second *P*, which has been preceded by two extrasystoles (x) of ventricular origin.

ventricles respond to the impulse coming from the auricle most of the time, but occasionally fail to do so, when the condition is called "dropped beat." The electrocardiogram records a P deflection but no ventricular complex, showing that the auricles have contracted at their usual rate but that the ventricles failed to respond to the stimulus coming from the sinoauricular node (Fig. 93).

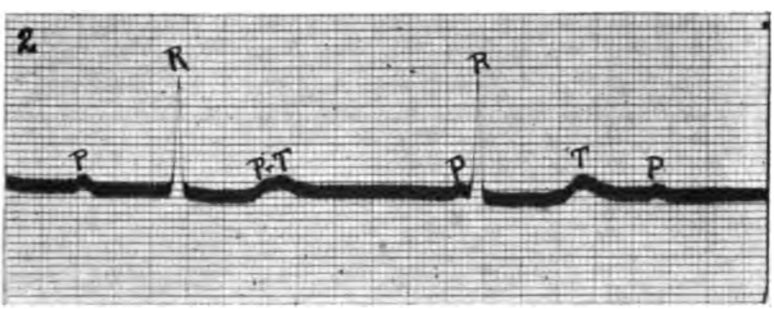


Fig. 94.—Complete dissociation. Note that the *P* wave spaces regularly and bears no definite relation to the *R* wave of the ventricular complex. Auricular rate 72. Ventricular rate 40.

Complete Dissociation.—In a simple case of complete dissociation the auricles beat independently of the ventricles; hence the P deflection of the electrocardiograms bears no relation to the ventricular complex (Q-R-S-T) (Fig. 94). The P deflections space regularly and are easily made out when they fall during diastole of the ventricle. Occasionally

the auricle will happen to contract during ventricular systole, causing a distortion of the ventricular complex by the superimposition of a P deflection. Except when this occurs the Q-R-S-T complex is the normal supraventricular type. The P deflections occur more frequently than the Q-R-S-T complex, showing that the auricles are beating more often than the ventricles. The auricular rate in the average case of complete heart-block is about 72, while the ventricular rate is much slower (35 to 40).

CHAPTER XXXI

CLINICAL APPLICATIONS OF CERTAIN PHYSIOLOGIC METHODS (Cont'd)

POLYSPHYGMOGRAMS

In taking polysphygmograms, the following technic is usually followed:

Venous Pulse Tracings.—Direct the observed person to lie down with his head slightly raised by a cushion and bent to the right side. Place the receiver (thistle funnel) over the jugular bulb on the right side of the neck. This lies immediately above the inner end of the clavicle. Bring the style of the recording tambour to write with a minimal amount of friction on the paper or drum. Since a venous pulse tracing can not be interpreted without a simultaneous tracing from an artery, now adjust the button of a receiving tambour over the radial artery and adjust the style of its recording tambour so as to write on the drum in the same perpendicular as the style of the venous tambour.

Tracings should be taken with the recording surface at a moderate speed. Before disturbing the relative positions of the writing points, allow them to make vertical marks (with recording surface stationary) at various parts of the tracings. These alignment marks permit of accurate comparisons between the curves. Repeat the above, using the carotid instead of the radial. A time tracing ($\frac{1}{5}$ sec.) should always be taken simultaneously. The polysphygmograph is shown in Fig. 95.

To interpret the venous curve, make a vertical mark on the arterial pulse tracing corresponding to the beginning of the pulse upstroke. If this is done on the radial pulse tracing, measure one-tenth of a second in front of it, and make a vertical mark to allow for the time lost in propagation of the pulse from the heart to the radial artery.

This line 3 (corrected in case of radial pulse) corresponds to the beginning of the sphygmic period of ventricular systole—i. e., to the opening of the semilunar valves. Measure the distance from it to the nearest vertical line that was made to indicate the relative position of the writing points. Then measure off the same distance in the venous tracing from the corresponding indicator line. This will fall at the beginning of the small wave (*c*), which is due to the bulging into the auricles of the closed auriculoventricular valves. (Fig. 96.)

The auricular wave (*a*) occurs one-fifth of a second in front of *c*, and may now be ascertained by measuring off this distance in front of *c*. This is line 1.

Now measure the distance on the radial pulse tracing from the beginning of the upstroke to the dicrotic notch. The distance between these is the sphygmic period (E).

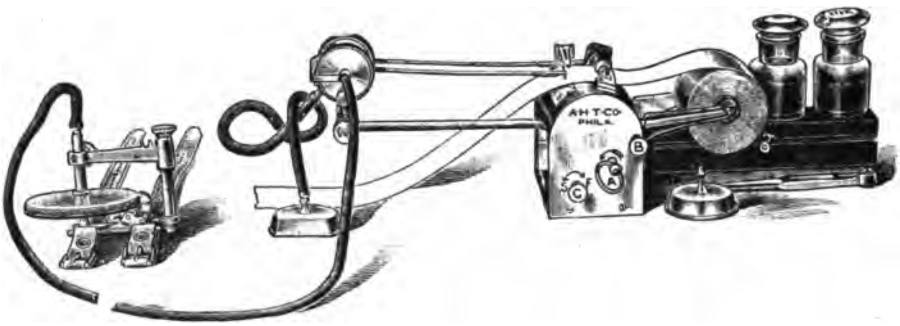


Fig. 95.—Polysphygmograph. This instrument records in ink on glazed paper two simultaneous tracings, i. e., radial pulse and one other, such as carotid, jugular, apex beat, etc., in addition to the time tracing. The ink tracings are both more convenient and permanent than smoked paper tracings. The clockwork operates at variable speeds, permitting the taking of protracted records at different speeds.

Measure off the same distance on the venous tracing from *c*. Line 5 will be found to fall just before a small wave (*v*), which is due to the sudden opening of the tricuspid valves. This practically coincides with the dicrotic notch on the radial pulse tracing. Sometimes a little wave

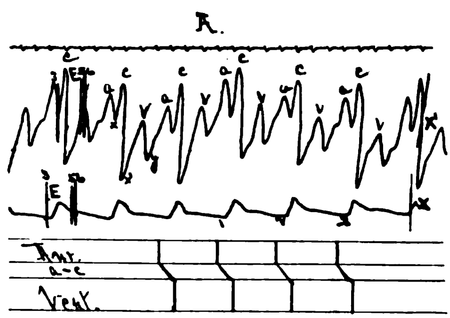


Fig. 96.—Normal jugular tracing. The spacing below shows the duration of the a-c interval. (From E. P. Carter.)

occurs on the upstroke of wave *v* just where line 5 falls. This coincides with the closure of the semilunar valves. The distance between it and wave *v* corresponds to the postsphygmic period.

The cause for the depression (marked *x*) following *c* will readily be understood by referring to the intraauricular curve (Fig. 97), to which, as already explained, the venous pulse tracing is qualitatively similar.

The rise in the curve following depression x is caused by the filling of the auricle with blood. This goes on until v, when the tricuspid valves open, allowing the blood to fall into the ventricle.

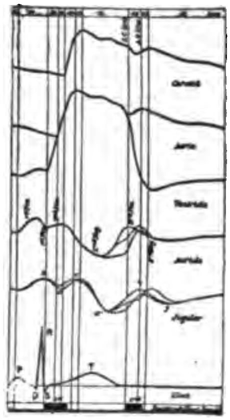


Fig. 97.—Reduced tracings from carotid, aorta, ventricle, auricle and jugular, to show the general relationships of the various waves. An electrocardiogram is also shown. Note that the jugular and auricular curves have the same contour, and that the depression (x) in them occurs during systole of the ventricles. (After Lewis.)

To interpret the cardiogram, adjust receiving tambours to the radial and open beat with both writing styles in the same perpendicular, and following the other directions described under "venous pulse" mark on the cardiogram: (See Fig. 98.)

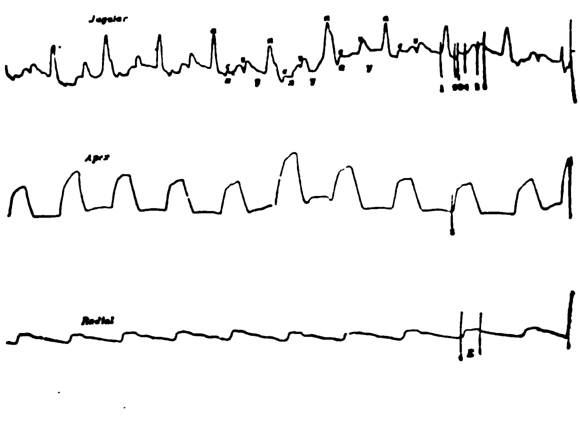


Fig. 98.—Polysphygmograms including jugular, apex and radial tracings. Line 4 on the radial tracing is first of all located. It is then transferred (by measurement from the alignment mark on the right edge of the tracing) to the jugular and 1/10 second subtracted from it, giving line 3. When this is similarly transferred to the apex tracing, it falls somewhere on the upstroke the beginning of which is line 2.

1. The beginning of the sphygmic period (E) (line 5).
2. The end of the sphygmic period (E) (line 5).
3. The auricular wave.
4. The beginning of ventricular systole (difference between 1 and 4 equals presphygmic interval).
5. The opening of auriculoventricular valves (lowest point in tracing).

Mark also on the tracing the exact moment at which the heart sounds are heard.

It is important to make certain that the button of the tambour is accurately over the apex beat, since otherwise a depressed or negative wave may be inscribed at ventricular systole.

Simultaneous Arterial Pulse Tracings.—The velocity of the transmission of the pulse wave is calculated by measuring the time between the systolic rise in the carotid and in the radial arteries, tracings of which are taken by applying one receiving tambour to the carotid artery and another to the radial artery.

Abnormal Pulses

The following is a brief description of the main character of abnormal pulses:

The Ventricular Pulse.—In this no "a" waves are present in the jugular tracing, the heart action being either regular or irregular. In the former case, the absence of the "a" waves may depend on: (1) overfilling of the right auricle, (2) increase in the heart rate, or (3) complete heart-block associated with auricular fibrillation. When the heart is irregular, the absence of the "a" waves signifies auricular fibrillation.

Delayed Conduction and Heart-block.—This causes a change in the time relationship of the "a" and "c" waves in the jugular curve. When the heart-block is of the first degree, the "a-c" interval merely becomes lengthened, but when it is of such degree that the normal impulse sometimes fails to be conveyed along the auriculoventricular bundle, isolated "a" waves can be detected. In the higher degrees of heart-block there are regularly recurring "a" waves having no constant time relationship to the "c" waves. For the purpose of exact analysis of the curves in suspected cases of delayed conduction, it is often advantageous to draw vertical lines below the tracing representing the beginning of auricular and ventricular systole. This has been done in the tracing reproduced in Fig. 99.

The line joining these two verticals indicates the conduction time or "a-c" interval. When it exceeds one-fifth of a second, there is delay in the conduction time.

A tracing showing a higher degree of heart-block is given in Fig. 100.

Sinus Arrhythmia.—In this condition the radial pulse is markedly irregular, but the "a," "c" and "v" waves of the jugular tracing occur with the usual time relationship to one another, and there is no delay in the "a"-"c" interval.

A.

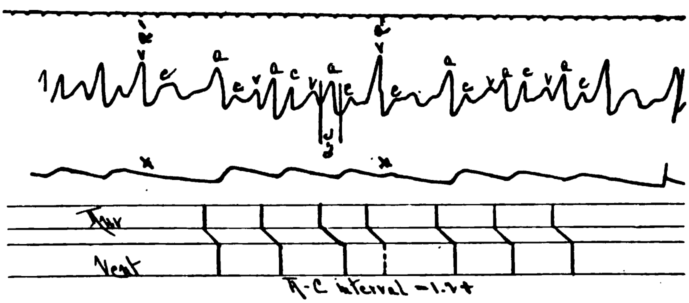


Fig. 99.—Delayed conduction time. First stage of heart-block. The *A-C* intervals measure more than 0.2 second. (From E. P. Carter.)

Sinus Bradycardia.—The beat originates at long intervals in the sinus; the "a-c" interval is normal, and the radial pulse very slow but practically regular.

Premature Beats.—These may be either ventricular or auricular in origin. In the former case the "a" waves on the jugular tracing space regularly throughout, but the "c" waves at the point of disturbance

B.

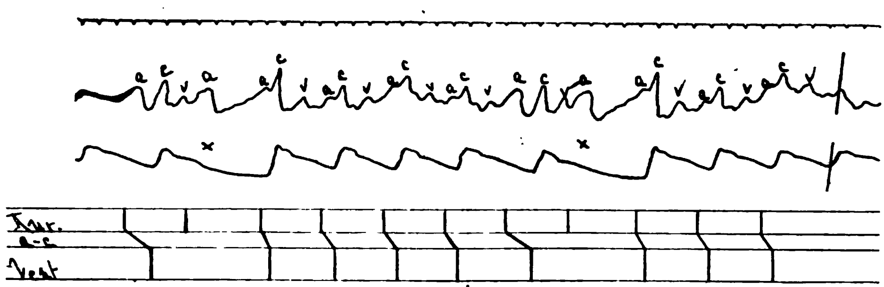


Fig. 100.—Dropped beats. Second stage of heart-block. (From E. P. Carter.)

coincide with the "a" waves, giving therefore a more pronounced wave. This is due to a premature contraction of the ventricle occurring about the time of the "a" wave, so that the latter finds the ventricle in a refractory state (see page 178). The premature contraction is therefore followed by a compensatory pause, which is evident on the tracing. An example of such a case is given in Fig. 101. In doubtful cases the exact

site of origin of the premature beats can be determined only by careful measurement of the distances between the various beats of the ventricle.

Whenever an irregularity repeats itself and the duration of one cycle of the arrhythmia accurately corresponds to another, the irregularity

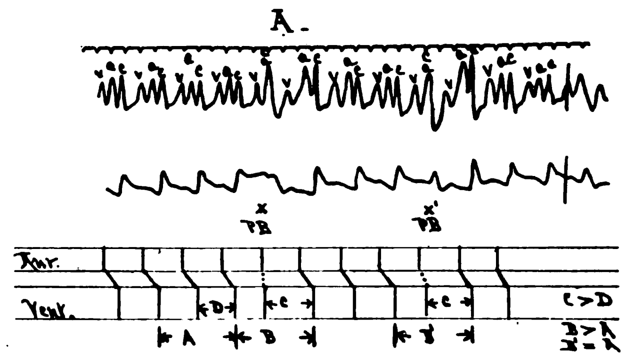


Fig. 101.—Premature beats (extrasystoles) ventricular in origin at *PB*. Compare the duration of the intervals marked *A* and *B'* with those marked *C* and *D*. (From E. P. Carter.)

may be due to: (1) premature auricular or ventricular contractions; (2) the occasional occurrence of dropped beats (a failure of ventricular response); or (3) a high degree of heart-block with a wide variation in the ventricular response. The important point to note here is that, no matter how irregular such a tracing may appear, if the irregularity repeats itself it can not be due to auricular fibrillation.

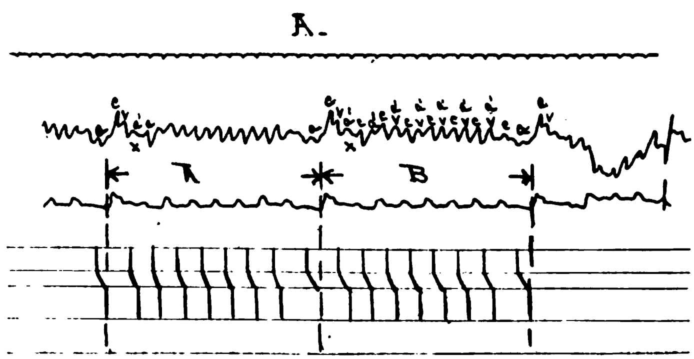


Fig. 102.—Paroxysmal tachycardia. The paroxysms start at *xx* following normal beats and lasting for seven beats. The clue to "a," which falls with "v" after the first premature contractions, is found in the initial beat of the new rhythm. (From E. P. Carter.)

Paroxysmal Tachycardia.—When the rate of a regular pulse is suddenly altered but the change in rate bears a simple ratio to the previous rhythm, the change may be due to (1) premature ventricular contractions which do not reach the radial, or (2) to the sudden development

of a two-in-one heart-block. When on the other hand, there is no exact ratio between the slow and the rapid rate, the change is due to the sudden appearance or disappearance of paroxysmal tachycardia. The paroxysms during which the auricle is beating very rapidly may last for a variable time, such attacks persisting off and on for hours or even days. The tracing in such a case is given in Fig. 102.

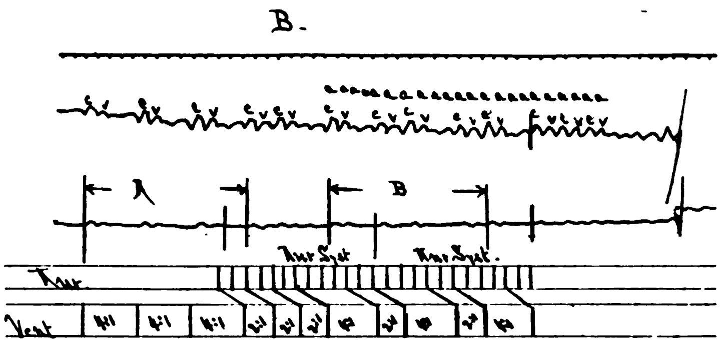


Fig. 103.—Auricular flutter. In this case the ventricular rate varied from 82 to 98 per minute. (From E. P. Carter.)

Auricular Flutter.—It is impossible to diagnose the not infrequent existence of this cardiac condition without the use of either the polysphygmogram or the electrocardiogram. The jugular curve may be of two types, one made up of rapid, more or less uniform waves, the other of waves that are paired with a constant time interval between the pairs.

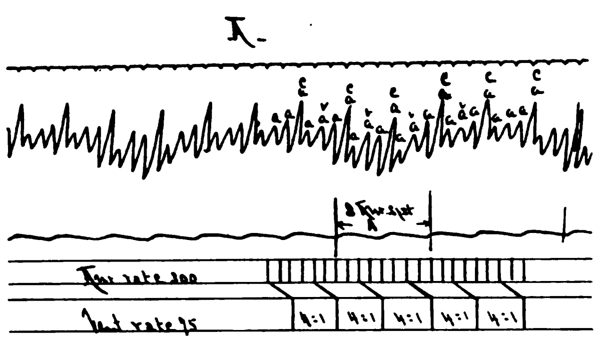


Fig. 104.—Auricular flutter. Note the relative rates of *A* and *V*, and also that the ventricular rate is regular. (From E. P. Carter.)

All of the frequent beats of the auricle do not reach the ventricle in this condition, so that the ratio between auricular and ventricular beats may be 1:3 or 1:4. The condition must therefore not be confused with heart-block, the main point of distinction being that in the latter condition the ventricular pulse is slow and the auricular about normal. The

radial pulse may be regular or irregular. The cause for the failure of every auricular beat to travel to the ventricle during auricular flutter is partly the refractory condition of the bundle, and partly the refractory phase of ventricular contraction. The bundle may be considered as a narrow bridge which will transmit the impulses across it only at a certain rate. If the impulses arrive too rapidly, only some of them can cross the bridge, and even of those that do cross, a limited number only will find the ventricle in a condition of excitability because of the refractory period (see page 178). Tracings showing auricular flutter are

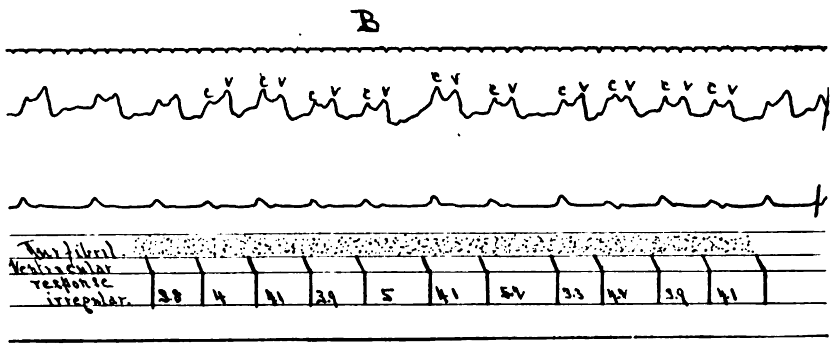


Fig. 105.—Auricular fibrillation. Note the absence of all "a" waves from the jugular tracing, the marked irregularity of the radial pulse, and the occurrence of "c" and "v" during the sphygmic period. (From E. P. Carter.)

given in Figs. 103 and 104. In one of them the radial pulse is regular; in the other, irregular.

Auricular Fibrillation.—The contractions of the auricle, as already explained, are entirely irregular, so that the jugular tracings show an entire absence of all "a" waves, the radial tracing being characterized by the complete absence of a dominant rhythm and by great variation in the length of the individual beats from one cardiac cycle to the next. This irregularity does not repeat itself, and the long pauses are not simple multiples of the shortest pause. Tracings from a case of auricular fibrillation are shown in Fig. 105.

CHAPTER XXXII

CLINICAL APPLICATIONS OF CERTAIN PHYSIOLOGIC METHODS (Cont'd)

THE MEASUREMENT OF THE MASS MOVEMENT OF THE BLOOD

Method.—The apparatus used for this purpose consists essentially of a vessel containing a known quantity (3,000 c.c.) of water and a thermometer from which a change of temperature of a hundredth of a degree centigrade can be read. In order to diminish as much as possible the loss of heat between the vessel and the outside air, the walls are double, the space between being stuffed with broken cork. The top of the vessel is closed with a thick cork plate, having suitable openings in it for the hand or foot and for the thermometer and a stirrer (feather) with which to keep the water in constant motion. The apparatus is called a *calorimeter.*

After the hand or foot has been in the calorimeter, with the water a few degrees below that of the body, for a certain time (ten minutes), the temperature of the water will of course become raised, and the degree to which this occurs, multiplied by the volume of the water in cubic centimeters, will give in calories the amount of heat dissipated. By the application of a very simple formula it is now an easy matter to calculate how much blood must have passed through the blood vessels of the part in order to give out the observed amount of heat; for, if we *divide the calories* by the *difference in temperature* between the inflowing and outflowing blood of the part, the result must indicate the *volume of blood,* in cubic centimeters, that has passed through it (since by definition a calorie equals volume multiplied by difference in temperature). It remains to explain the equation by which the results are arrived at. If Q equals the amount of blood, H the calories of heat given out to the calorimeter, T the temperature of the arterial blood and T' the temperature of the venous blood, then we have the equation: $Q = \frac{H^*}{T - T'}$. It has been shown

*For the determination of H we must multiply the cubic centimeters of water *plus* the water equivalent of the hand and calorimeter (because both of these will absorb some heat) by the difference in temperature *plus* the self-cooling of the calorimeter (because some heat is lost to the air during the observation). The water equivalent of the hand is equal to its volume multiplied by 0.8; that of the calorimeter must be determined for each instrument and is usually about 100 c.c. The self-cooling of the calorimeter is determined by observing the fall in temperature for a period equal to that of the actual observation without the hand in the calorimeter.

by Stewart that T may be taken as the same as that of the mouth, or 0.5° C. below that of the rectum, and T' as the average temperature of the water in the calorimeter during the observation. To allow for the specific heat of blood, the result is multiplied by $\frac{10}{9}$, the reciprocal of the specific heat of blood.

Theoretically, then, the method is very simple, and there are no unusual technical difficulties in applying it. The only special precaution is that the air surrounding the calorimeter should be kept fairly constant in temperature, so that we may be enabled to allow in our calculations for the loss of heat from the calorimeter itself, this value being obtained by observing the change of temperature in the calorimeter for a certain period of time after the hand has been removed from it.

The Normal Flow

The results are calculated on the basis of grams of blood flowing through 100 c.c. of tissue in one minute. The volume of the hand or foot is ascertained by placing it in water contained in a small-sized irrigation can, the tube of which is connected with a burette. The height to which the water rises in the burette is noted, and after withdrawing the hand, water is added from a graduate to the irrigation can until the same height is reached on the burette. The number of cubic centimeters required gives the volume of the hand. In a normal, healthy individual the average flow in the hand is from 12 to 13 gm. for the right hand, and about half a gram less for the left. This difference between the two hands corresponds, of course, with their relative degree of development. The average foot flow is much less, and varies according to whether the patient is sitting up or lying down while the measurement is being made. In a normal individual, while lying down, it was 5.11 gm. in the right foot and 5.23 gm. in the left, per 100 c.c. of foot; but only 2.96 gm. for the right and 4.1 gm. for the left foot, while sitting up. The results have been found to undergo only a slight variation from month to month in a given healthy individual, provided the air temperature during the different observations is the same and the person has been some time in the room before the observations are begun. This precaution is especially important if he is a dispensary patient and has been in the open air with bare hands. The flow varies in different individuals both with regard to absolute amount and the ratio between the two hands or feet. When the total flow in the hands is compared with that in the feet, a ratio of about 3 to 1 is usually obtained.

The Physiologic Causes for Variations in Bloodflow.—As above indicated, the most marked of these is probably the *temperature of the room.* The

temperature of the water in the calorimeter has likewise a great influence, and for the comparison of different cases it is always important that the room and calorimeter temperatures be stated alongside the results. *Muscular contractions,* produced by moving the fingers in the calorimeter, cause a marked increase in flow, accompanied by a diminished flow in the hand that was at rest. A great diminution of flow results from *constriction of the arm* of sufficient degree to obstruct the venous circulation; and when the constriction, as that caused by a blood pressure armlet, is increased to between the systolic and diastolic pressures, extremely little blood flows through the hand.

By immersing the opposite hand or foot in hot or cold water, the bloodflow through the observed hand increases or decreases, respectively. The change may be of a temporary character, or it may persist throughout the whole period of immersion of the hand. These reactions are due to a *vascular reflex,* and observations of its sensitiveness are of value in the study of the effects of lesions either of the nerve or of the nerve centers concerned in vascular reflexes.

Clinical Conditions which Affect the Bloodflow

Even in cases where there is plenty of other evidence of curtailment of flow, the measurement may be of importance either for detecting an alteration in the vascular reflex or, by comparison of the two hands, for demonstrating the relative degree of alteration in flow. In *acute inflammatory conditions* affecting one hand, there is an increase in flow on the affected side accompained by a marked curtailment on the other side. This indicates that an increased flow in the infected area is accompanied by a reflex vasoconstriction elsewhere, particularly in the symmetrically placed part of the opposite side of the body. In cases of nonbacterial inflammation of the hand, as in gout, no sign of vasoconstriction may be observed.

There are many clinical conditions in which Stewart's method reveals an alteration in bloodflow that would be unsuspected by the use of ordinary clinical methods. It is for the investigation of these that the method is of greatest value. The most important findings are as follows:

Anemia.—The bloodflow in the hand may be much less than normal in pernicious anemia and secondary anemia, and distinctly curtailed in chlorosis. Since the minute volume of the heart is also increased in these conditions, the vasoconstriction at the periphery will assist in compelling more blood to pass through the lungs, so as to make up for deficiency of blood.

Fever.—Since changes in the cutaneous circulation probably constitute the chief factor in the derangement of the temperature-regulating mechanism in fever (cf. page 723), it is evidently of great advantage to be able to measure such changes quantitatively. This has been done by Stewart in several cases of typhoid fever and in one case of pneumonia. In general it was found that the flow in the feet never exceeded the normal flow, and was usually much below it. This tendency to vasoconstriction seems to be carried into convalescence. For practical reasons the handflow has not been so extensively studied. This hyperexcitability of the vasoconstrictor mechanism at the periphery is most naturally interpreted as a defensive reaction of the organism by which an increased supply of blood is imported to those internal organs which bear the brunt of the infection. When we consider that in spite of this constriction of the periphery the blood pressure is low and the pulse dicrotic, we must conclude that there is considerable dilatation of other vascular parts, especially the splanchnic area. A very practical application of these facts presents itself in considering the rationale of the cold-bath treatment for fever. If, for example, we conclude that the cutaneous constriction is in the interests of an increase in the bloodflow to the organ on which the stress of the infection falls, it will evidently be more rational to lower the temperature by methods which will not diminish, and may even increase, the cutaneous constriction than to do so by causing the vessels to dilate. In other words, the use of antipyretics seems to be contraindicated, since they diminish the body temperature by causing vasodilatation at the periphery with a consequent withdrawal of blood from the seat of infection.

Cardiovascular Diseases.—In cardiac cases the handflow is far more apt to be markedly deficient where there is evidence of serious impairment of the myocardium than in cases where a gross valvular lesion exists but the heart action is strong and orderly. This indicates that it is more serious for the force of the heart pump to be interfered with than for its valves, particularly the mitral, to be leaky. Even where there is considerable venous engorgement, the flow may be little diminished. In untreated cases of auricular fibrillation the bloodflow is subnormal. After the administration of digitalis the bloodflow in such cases is often promptly and decidedly increased.

As would be expected, *arteriosclerosis* is associated with a small bloodflow, and the vasomotor reflexes are weaker than in normal persons.

In *aortic aneurism,* when the aneurism is of such a size as to cause pressure on the subclavian artery or vein, there is a diminution in flow of the corresponding hand, but aortic aneurism itself, although it may

cause great changes in the character of the pulse beat, does not decidedly affect the mass movement of the blood. In aneurism of the subclavian artery, the bloodflow may be much greater in the corresponding than in the opposite hand, even though the amplitude of the pulse is very obviously diminished and the difference between the systolic and diastolic pressures (the pressure pulse) is much less on the affected than on the normal side. By ordinary clinical measurements, therefore, false estimates of bloodflow are quite likely to be made. These results are no doubt owing partly to vasodilatation brought about by pressure of the aneurism on the brachial plexus and partly to the lower resistance to the flow of blood into the dilated subclavian.

In *Raynaud's disease,* as would be expected, the flow is small, the diminution being more or less proportional to the duration of the disease. The contralateral vasomotor reaction to cold is also peculiarly intense.

In *diabetic gangrene* of the feet there is a very subnormal flow in both the hands and the feet. The vasomotor reflexes are also feeble.

It is sometimes difficult to tell whether an observed curtailment of flow is a nervous (reflex) effect or is due to some mechanical interference. There are two ways by which the exact cause may be diagnosed: (1) by observing the flow from day to day; if it remains unchanged, any alteration must be dependent on mechanical causes; (2) by observing the change in flow brought about by altering the temperature of the room or calorimeter and seeing whether the ratio between the two hands remains unchanged or becomes altered. If the latter occurs, the inequality in flow must be due to nervous causes.

Diseases of the Nervous System.—The effect of *neuritis* on the flow varies with the duration of the disease. In cases of early peripheral unilateral neuritis there may be an increase of flow altering the ratio between the two hands with the greater flow on the diseased side. In neuritis of long standing the flow is cut down, the greater flow occurring on the healthy side. The changes here are probably due to anatomic alterations in the lumen of the tube, perhaps a thickening of the intima. In *motor-neuron disease* without any involvement of the sensory skin nerves the flow seems to remain normal and the reflexes to be well-marked. This indicates that involvement of the motor nerves does not interfere with bloodflow to anything like the same degree as involvement of the skin nerves.

Hemiplegia.—A deficiency of bloodflow of the paralyzed side is usually observed, and the vasomotor reflexes are altered, the most usual change being that vasoconstriction is more easily produced than vasodilatation.

In some cases an abnormal tendency to vasoconstriction is a conspicuous feature.

Tabes Dorsalis.—The flow is distinctly diminished, especially in the feet, although also in the hands, and the vasomotor reflexes are feeble. Sometimes there is inequality in the flow of the two hands, which however need not necessarily indicate a unilateral lesion of the cord in the cervical region.

CHAPTER XXXIII

SHOCK

Shock may be due to a variety of causes. In general it may be described as a condition in which there is more or less paralysis of the sensory and motor portions of the reflex arc, along with profound disturbances in the circulatory system, subnormal temperature, frequent and shallow respiration, and more or less unconsciousness. Certain of these symptoms may be considered as primary and others as secondary, an important step in the investigation of this difficult and important problem being to distinguish between the two groups. Before attempting to do this, however, it will be profitable to differentiate as sharply as possible the various conditions in which one or another of the many varieties of shock is said to occur.

The following varieties of shock have been described:

1. **Gravity Shock.**—This is caused by the stagnation of blood in the splanchnic vessels and the consequent inadequate filling of the heart in diastole. It occurs, when the erect position is assumed, in animals in which the mechanism which ordinarily compensates for the tendency of gravity to make the blood flow to the dependent parts is inadequate. Thus, when a domesticated rabbit with a large pendulous abdomen is held in the vertical tail-down position for any length of time, the animal gradually passes into a shocked condition and may die in a short time (20 to 30 minutes). Observation of the blood vessels of the ear or a record of arterial blood pressure will show that the cause of shock in this case has been a great curtailment of the blood supply to the upper part of the body, and therefore to the nerve centers (Fig. 244). The shock is entirely dependent upon the laxity of the abdominal musculature, for if a binder is applied to the abdomen, or if the experiment is performed on a rabbit whose abdominal musculature is in good condition, gravity shock does not develop. Nor can fatal gravity shock be produced in a dog, although in a deeply anesthetized animal a marked fall in arterial blood pressure occurs when the vertical tail-down position is assumed. In man, in whom compensation for the erect posture is highly developed, shock from gravity occurs only when there has been some other considerable upset in the circulatory mechanism (see also page 245).

2. **Hemorrhagic Shock.**—Free hemorrhage produces a typical condition of shock, but the extent to which different individuals react to the same degree of hemorrhage varies considerably. The essential factor in the production of hemorrhagic shock is of course similar to that of gravity shock—namely, a deficient diastolic filling of the heart with blood. Details concerning the effect of hemorrhage will be found elsewhere (page 135).

3. **Anesthetic Shock.**—So far as blood-pressure reflexes are concerned, an animal can be kept in a perfect condition when ether is administered in just sufficient amount to produce light anesthesia. When larger quantities of ether are employed, a typical condition of shock is almost certain to supervene after a time. In such instances the arterial blood pressure remains low and can not be restored even after an hour or two of artificial respiration. There is, however, a difference between ether shock and the variety which we shall discuss later under the title of surgical shock: in the former, removal of the anesthetic causes all reflexes to return, whereas in surgical shock most of these are subnormal. The danger of anesthetic shock has been considerably diminished in the clinic by the more careful administration of ether or by the use of other anesthetics, such as nitrous oxide gas. A condition closely simulating shock may also be induced in the earlier stages of the administration of anesthetics when these are badly given, but paralysis of the heart or of the respiratory center is a usual contributory cause.

4. **Spinal Shock.**—Spinal shock is produced by section of the spinal cord, but it is to be carefully distinguished from all other forms of shock because of its local character, as it affects only those parts of the body which lie below the level of the lesion in the cord. Above this level the animal may be in a perfectly normal condition, except in cases where the section has been at so high a level that it has severed the vasoconstrictor pathway and thereby produced a fall in blood pressure from vasodilatation. Even when this has happened the part of the animal anterior to the spinal lesion is by no means in a condition of shock. Thus, Sherrington observed in a monkey whose spinal cord had been cut far forward that, although the posterior part of the body was in profound spinal shock and the blood pressure very low, the animal amused himself by catching flies with his hands. A sufficient description of the condition of spinal shock has been given elsewhere, but here it may be noted that it consists essentially in a paralysis involving at first all of the reflex mechanisms, including the control of the sphincters, in the part of the cord posterior to the section. In the course of a few days or weeks, according to the position of the animal in the scale of development, the reflexes gradually return, until ultimately in a couple of months—in a

dog, for example—they may all have reappeared. The cause of this shock is no doubt the sudden interruption of the nervous pathways which reflex action ordinarily takes in the higher animals (see page 803).

5. **Nervous Shock; "Shell Shock."**—Considerable attention has been paid to the nervous shock that has frequently been observed in men who have been subjected to the harrowing sights and the constant noise and nerve strain incurred in modern warfare. The symptoms may appear suddenly at the front or they may develop in men who have comported themselves in an apparently normal manner until removed to the rear, when they pass into a condition more or less simulating that of shock. Severe conditions may also result to soldiers from injuries which in normal individuals would not in themselves be sufficient to produce surgical shock. The characteristic symptoms in such cases are entirely different from those of other forms of shock, and, as has been shown by Elliot-Smith and T. H. Pear,[25] the condition must be treated from the neurologic or psychopathic point of view.

6. **Surgical Shock.**—It is this variety that is usually referred to when one speaks of shock. It may be produced either by severe mechanical injury to a healthy person or by extensive manipulation and rough handling on the operating table. It is common in trench warfare, being therefore an important variety of "shell shock," which term must be used only in a general sense. However produced, the symptoms of surgical shock are very much the same. The patient lies in a quiet, apathetic condition, caring little for what is going on around him, and answering questions only when repeatedly and importunately questioned. His skin, lips and gums are very pale and more or less cyanotic; the skin feels cold and is moist with sweat; the reflexes are greatly diminished, and it is usually only after applying a very painful stimulus that any movement of defense is elicited or resentment is shown on the part of the patient. The postural reflexes are also abolished, so that if a limb is lifted it falls back limp and toneless. The pulse at the wrist is very rapid, thin and almost imperceptible, and the arterial blood pressure is abnormally low. The respirations are frequent and shallow. The rectal temperature is 1° C. or more below normal. The pupils are dilated and react slowly to light. When he can be induced to speak, the patient's voice is hoarse, and he complains of cold and numbness in the extremities. The symptoms are not unlike those of cholera.

Experimental Investigations of Shock

For inducing shock experimentally, two general methods have been employed: either rough manipulation of the abdominal viscera, or re-

peated stimulation of large afferent nerves. Since the experiments are usually performed on anesthetized animals, the effect of the anesthetic has to be discounted in experimental work on the causes of shock.

The first step in such an investigation is naturally to classify the symptoms into primary and secondary, for on the success of the classification must depend the outcome of further investigation into the problem.

The earlier investigators were naturally attracted to *the pronounced fall in blood pressure* as the primary cause of shock. It is true that a pronounced lowering will ultimately produce symptoms that are not unlike those of shock, but on the other hand it can readily be shown that this is a result only—a symptom and not a cause of the condition. It was believed by Crile that the fall in blood pressure depended on a universal dilatation of the blood vessels caused by exhaustion of the tone of the vasoconstrictor center. It has been clearly shown, however, that the tone of this center is practically normal in shock, and that the arterioles are maintained not in a dilated but in a contracted state, indicating clearly therefore that the low blood pressure must be dependent upon inadequate output of blood from the heart. The evidence for this conclusion is as follows: (1) W. T. Porter[26] and his collaborators have shown that both pressor and depressor reflexes are perfectly normal in a rabbit that is in a condition of extreme shock. It is particularly important that depressor effects were still obtained, since this indicates that tonic activity of the center must still have been present. (2) The blood vessels in a shocked animal are in a contracted state. On opening a vessel and observing the outflow of blood, an increase occurs when the nerve to the blood vessel is cut. (3) This same fact has been shown by Seelig and Joseph,[27] who cut the vasomotor nerve proceeding to the vessels of one ear of a white rabbit and thus caused a local paralytic dilatation of the vessels. Intense shock was then produced in the animal in the usual way, after which the blood pressure in the anterior part of the animal was suddenly raised by applying a clamp to the abdominal aorta just below the diaphragm. This increased blood pressure caused the vessels of the denervated ear to become engorged with blood, but not those of the opposite normal ear, which retained their tone (Fig. 106). (4) The volume of blood expelled by the ventricles has been shown by Henderson[28] to be distinctly diminished in the early stages of shock, the lack of pronounced fall in blood pressure indicating that there must be a compensatory constriction of the arterioles. Lastly (5), it has been found by Morrison and Hooker[29] that the outflow of blood from the perfused organs of a shocked animal is less than that from the

Fig. 106.—Illustration showing the appearance of the blood vessels in the ears of a rabbit "in a state of deep shock." The marked vasoconstriction is very plain in the left ear, the vessels of the right ear being dilated because the cervical sympathetic, which carries the constrictor fibers, has been cut. (From Seelig and Joseph.)

same organs under normal conditions. Furthermore, severing of the nerve of such an organ causes an increased outflow.

To these various pieces of evidence of a constricted condition of at least certain of the vessels in shock, may be added the less direct evidence furnished by the pallor of the shocked patient and the indications that the sympathetic nervous system, instead of being paralyzed, is in an excited state, as shown by the sweating and the dilated pupils.

Furthermore, we know from the experiments of Pike, Guthrie and Stewart[30] on the resuscitation of the nerve centers after interference with the circulation to the brain, that the vasomotor center is remarkably resistant to anemia. It can withstand this condition without losing its tone or reflex activity better than any of the other cardinal centers.

Those who have maintained that a deficiency in the tone of the vasoconstrictor and other nerve centers is responsible for shock have based their evidence partly on histologic examination of nerve cells of shocked animals, it being assumed that the chromatolysis shown by these cells indicates an exhausted condition. The assumption is, however, entirely unwarranted, and no regard is given to the well-established fact that similar histologic changes may be produced by other conditions. It is certainly safe to conclude that the changes in the nerve cells in shock are the result and not the cause of the low blood pressure of this condition.

Since the fall in arterial blood pressure occurs with contracted arterioles, it must be dependent on *a diminished discharge of blood from the heart*. Interference with the heart action itself (independently of the blood carried to this organ), or a deficiency in the filling of the ventricles during diastole,—that is, a stasis of blood in the venous or capillary areas,—are the possible causes for the diminished output. The possibility that the heart action itself has been interfered with, as by paralysis of the vagus mechanism, causing a rapid beating of the heart, has been shown to be untenable by various experiments. After stimulating the central end of an uncut vagus nerve in the neck in shock, the reflex vagus mechanism is still operative. Furthermore, when the arterial blood pressure is artificially raised, either by epinephrine injection or by cerebral compression, the heart promptly responds to the increased blood pressure by contracting more slowly and vigorously. Evidently, therefore, as the cardiac mechanism itself is normal, the deficient discharge of blood must be dependent upon *improper diastolic filling*. After this condition has set in, it becomes progressively worse because of weakening of the heart muscles consequent upon the failing blood supply through the coronary vessels.

The question therefore narrows itself down to the cause of the ineffi-

cient return of venous blood to the heart. In the first place, let us see whether shock can be produced experimentally in animals by mechanical interference with the bloodflow in the vena cava. That such is the case was shown by H. H. Janeway and Jackson,[31] who found that mechanical obstruction of the inferior vena cava for a short time was followed by the usual signs of shock. Such interference with the venous return to the heart may also be produced by excessive movements of the thorax as a result of artificial respiration. That this in itself may cause shock is known to all experimental investigators on the subject, although the interpretation has not always been that which is given above. Yandell Henderson[32] thought that the excessive ventilation caused a blowing off of carbon dioxide from the blood (see page 293), thus producing a low tension of this gas in the blood (acapnia), which he believed to be the responsible factor.

As in gravity shock, so in surgical shock, *stagnation of blood in the splanchnic area is common;* the animal bleeds into his own (splanchnic) blood vessels (capillaries and venules), because these have lost their tone. As we have noted above, one of the most certain ways of producing shock is by exposure and rough handling of the abdominal viscera. It is therefore of importance to study the effects that can be noted on the blood vessels of this area under such conditions. When the viscera are first exposed to air, there may be a short period during which vasoconstriction is evident. This is soon followed by a dilatation of the arterioles in the exposed area, causing the capillaries and veins to become markedly distended as during the first stage of inflammation. This accumulation of blood in the mesenteric veins has been shown by Morrison and Hooker to cause an increase in the weight of an isolated loop of intestine as an animal passes into a state of shock.

Splanchnic engorgement alone does not, however, suffice to explain all the loss of blood, and we are driven to conclude that the capillaries of the tissues outside the abdomen must entrap much of it. As a matter of fact, Cannon, and others, have found that concentration of the blood occurs in these capillaries as indicated by comparisons of the corpuscles and hemoglobin in blood drawn from veins and from capillaries. Normally the values are equal; in shock the capillary blood is much concentrated.

In so far as the circulatory disturbances are concerned, we may therefore sum up the conditions occurring in shock as follows: The blood accumulates in the veins and capillaries—that is, in a part of the vascular system that is beyond vasomotor control. The consequent withdrawal of this blood from the circulation produces a diminution of the bloodflow in the vena cava and consequently an inadequate filling of the

heart. The consequent curtailment in the systolic discharge does not, however, at first cause any marked fall of arterial blood pressure because of a reciprocal constriction of the peripheral arterioles of the body. Meanwhile, however, the stagnation of blood in the capillary areas is progressively increasing, so that less and less blood remains available for the systemic circulation. Consequently, after a while, in spite of the arterial constriction, the blood pressure falls to the dangerous shock level, and the secondary symptoms of fall in temperature, dulling of the reflexes, etc., supervene. Increasing viscidity of the blood also retards its flow.

The fundamental question in the pathogenesis of shock concerns therefore *the cause of the stagnation of circulatory fluid in the capillaries and venules.* Two hypotheses have been offered, one being that the stimulation of afferent nerve fibers to the respiratory center causes excessive alveolar ventilation with a consequent washing out of carbon dioxide from the blood (acapnia), which causes a veno-capillary atonia, and the other, that a bombardment of the vasoconstrictor and other nerve centers by afferent impulses brings these centers into a condition of exhaustion, which is the essential cause of shock. The acapnia hypothesis may be at once dismissed, since, on the one hand, it has been shown that in typical shock there is no deficiency of carbon dioxide in the venous blood (Short),[33] and on the other hand, conditions of shock are often produced without excessive breathing.

Nor is there any evidence to support the view that shock is caused by fatigue of the cardinal centers as a result of excessive sensory stimulation. In the first place, it has been shown by Mann[34] that during handling of the abdominal viscera the nervous impulses transmitted up the spinal cord are much less marked than those transmitted when the central ends of sensory nerves are stimulated by operative processes on the limbs and joints, although shock is much more readily produced by the former procedure. The method employed by Mann for detecting the existence of these afferent impulses was that of Forbes and Miller, in which electrodes are placed on the brain stem in decerebrate animals, and the current of action which accompanies the passage of nerve impulses registered by a string galvanometer. Although this method is simple and direct in principle, it has been found by Mann to require great care in practice because of the fact that the slightest movement of the head end of the animal produces deflections of the galvanometer. If the further results of this investigation should show, as the early ones have done, that shock may be produced in an animal without any observed deflection of the galvanometer, it will disprove once and for all

the theory that shock is dependent upon an impairment of a higher nerve mechanism as a result of overstimulation by afferent impulses.

Cannon[35] has recently suggested that the engorgement of the splanchnic blood vessels may be the result of a constriction of the portal radicles in the liver, which dams back the blood in the portal circulation. He points out that these radicles have vasoconstrictor nerve fibers, as evidenced by the fact that the rate of flow of fluid through the perfused liver decreases during asphyxia, as well as when the hepatic nerve plexus is stimulated electrically or when epinephrine is injected into the portal vein. He argues that, since the blood vessels in other areas of the body are constricted in shock, so also will be those of the liver, with the result that the blood of the portal vein, in which ordinarily there is a very low blood pressure (10 mm.Hg), will become dammed back in these veins and therefore removed from the systemic circulation. It does not seem to the writer, however, that this explanation is likely to be the correct one, for, although it is true that vasoconstrictor influences have been shown to exist in the hepatic radicles of the portal vein, yet, since it is only under special experimental conditions that this can be done, they must be very feeble in nature. As we have seen elsewhere, portal vasoconstriction can not be demonstrated by stimulation of the hepatic plexus with stimuli which are sufficient to produce marked constriction of the hepatic artery radicles (see page 255).

The engorgement of the splanchnic capillaries and venules is much more likely to be dependent upon *local influences acting on the vessels themselves.* When shock is produced by manipulation of the abdominal viscera, it is easy to see how this local disturbance might be set up. When shock is caused in other ways, as by violent stimulation of sensory nerves, the atony of the splanchnic vessels is not so easily accounted for unless we assume that it is a type of abnormal reciprocal vascular innervation. For example, when stimuli are applied locally to sensory surfaces under ordinary conditions, a distribution of the blood of the body takes place, more being sent to the irritated region and less to other parts of the body (see page 238). During the sensory stimulation preceding shock, it is conceivable that this reciprocal innervation acts in a faulty manner, causing at first a dilatation of the splanchnic arterioles and thus allowing more blood to enter the splanchnic capillaries and venules, which being possessed of little tone are incapable of responding by increased tonicity, so that they become overdistended and the blood in them stagnates.

In any case there is no doubt that the initial change is the stagnation of blood in these vessels, and when once such stagnation has occurred, the process goes on spontaneously probably on account of the accumula-

tion in the stagnant blood of incompletely oxidized metabolic products, which raise the hydrogen-ion concentration of the blood, and produce a further relaxation of the muscle fibers in the vessel walls. That acid has such an effect is well known (page 937). Dilatation or atonicity thus progressively increases and is meanwhile further encouraged by the deprivation of oxygen, for it has been shown that isolated artery strips do not exhibit their usual tonicity when deprived of oxygen.

Treatment

Whatever may be the cause of the atony of the capillaries and venules in shock, the existence of this condition indicates that the most hopeful line of treatment is to cause the vessels to reacquire their tone. It will be remembered that in gravity shock in a rabbit recovery may be accomplished by the application of a tight binder to the abdomen, or by placing the animal in a head-down position. Such measures applied in the case of man have not, however, been found of much value. Pressure thus applied is evidently not brought to bear sufficiently on the atonic vessels. Cannon has therefore made the interesting suggestion that a hopeful procedure may consist in injecting directly into the abdomen a saline solution containing pituitrin, a hormone which, it will be remembered, acts directly on involuntary muscle fiber.

Two other methods have been advocated for the treatment of shock—namely, saline or blood transfusion and injection of epinephrine; but neither method has proved of practical value. Epinephrine injections do indeed temporarily raise the arterial blood pressure, but the subsequent condition of shock is possibly worse than that originally present. After the injection of blood or saline solution containing gelatin or mucilage, the blood pressure, although temporarily raised, very quickly falls again. In this regard surgical shock differs from the shock following severe hemorrhage, in which, as explained elsewhere, recovery of the blood pressure as well as of the general condition of the animal can be accomplished by transfusion with blood or with saline solution containing mucilage or gelatin. This would indicate that there is some essential difference between these two forms of shock (see page 140). The only treatment of avail appears to be to keep the patient warm and to remove causes of excessive pain.

Causes of Secondary Symptoms

It remains to consider the cause of some of the secondary conditions developing in shock—namely, the *disturbances in sensation and motion and the fall in body temperature.* All of these are undoubtedly depend-

ent upon the low arterial blood pressure, although some authors have suggested that the loss of sensation may be dependent upon an increased resistance or block at the synapses of the receptor neurons (page 803). This suggestion depends on the fact, demonstrated by Sherrington, that repeated stimulation of the receptors of a reflex arc produces fatigue of that particular reflex, and that this fatigue must be resident in the synapsis and not in the motor neuron, since the same motor neuron that participated in the fatigue can still be called into activity by afferent stimuli transmitted to its nerve cell through other sensory pathways (see page 825). It is thought that in shock the frequent afferent stimulation produces synaptic fatigue and therefore dulls the sensory responses of the animal. The researches of Mann above referred to, in which he shows that shock may occur without any demonstrable afferent stimuli in the brain stem, would seem, however, to negative the above hypothesis.

The raised threshold of sensory stimulation is no doubt an effect of the low blood pressure. It has been shown, for example, by E. L. Porter[36] that when the arterial blood pressure is maintained at a uniform level, the threshold stimulus for spinal cord reflexes remains practically uniform, but becomes promptly increased when the arterial blood pressure is made to fall. Why a lower blood pressure should have this effect is, however, difficult to understand in the light of the researches of Stewart and his coworkers, who, as remarked above, found that the cells of the central nervous system may endure total anemia for many minutes and still recover their physiologic condition. It may be, however, that the low blood pressure affects the conductivity of the synapsis.

The muscular weakness is probably also dependent on low blood pressure, for it has been found in animals that, when the arterial blood pressure is lowered to about 90 mm. Hg, the muscles contract much less efficiently than ordinarily. The fall in body temperature is dependent upon the muscular inefficiency.

In conclusion, it should be pointed out that W. T. Porter, in the investigation of acute shock met with at the front, has found that, in many cases at least, the circulatory disturbance is due to a condition of fat embolism. The fat is derived from the marrow of long bones, such as the femur, by injuries which smash the bones. Porter's observations are at least very suggestive.

CIRCULATION REFERENCES

(Monographs)

Wiggers, C. J.: The Circulation in Health and Disease, Philadelphia, 1915.
Mackenzie, J.: Diseases of the Heart, Oxford Medical Publishers, ed. 2, 1910.
Lewis, Thomas: Mechanism of the Heart Beat, 1911, Shaw & Son, Fetter Lane, London.

Lewis, Thomas: Harvey Lectures, 1913-1914, J. B. Lippincott Co.
Lewis, Thomas: Clinical Disorders of the Heart Beat, P. B. Hoeber, New York, 1912.
Hill, Leonard: The Mechanism of the Circulation of the Blood, in Schäfer's Physiology, ii, 1900. Young J. Pentland.
Gaskell, W. H.: The Contraction of Cardiac Muscle, in Schäfer's Physiology, ii, 1900, Young J. Pentland.
Flack, M.: Further Advances in Physiology, 1909. Ed. by Leonard Hill, E. Arnold, London.
Porter, W. T.: American Text Book of Physiology, W. B. Saunders Co., 1900.

(Original Papers)

[1]MacWilliam, J. A., *et al.*: Heart, 1913, iv, 393; ibid., 1914, v, 153; Brit. Med. Journal, Nov., 1914; VII Internat. Congress of Medicine, London, 1913, Sec. II, Physiology.
[2]Hill, Leonard, F. R. S., *et al.*: Proc. Roy. Soc., 1914, *B*, lxxxvii, 344; ibid., 1915, *B*, lxxxviii, 508 and 516.
[3]Erlanger, J.: Am. Jour. Physiol., 1916, xxxix, 401; ibid., 1916, xl, 82.
[4]Downs, A. W.: Am. Jour. Physiol., 1916, xl, 522.
[5]Bayliss, W. M.: Proc. Roy. Soc., 1916, lxxxix, *B*, 380.
[6]Knowlton, F. P.: Jour. Physiol., 1911, xliii, 219.
[7]Milroy, T. H.: Jour. Physiol., 1917, li, 259.
[8]Eyster and Meek: Heart, 1914, v, 119; ibid., 194, v, 137; Am. Jour. Physiol., 1914, xxxiv, 368.
[9]Porter, W. T.: Art. on Circulation in an American Textbook of Physiology, W. B. Saunders Co., 1900.
[10]Brodie, T. G.: Proc. Physiol. Soc., 1905, Jour. Physiol., 1905, xxxii.
[11]Stewart, G. N.: Heart, 1911, iii, 33.
[12]Garrey, W.: Am. Jour. Physiol., 1912, xxx, 451.
[13]Mines, G. R.: Jour. Physiol., 1913, xlvi, 188.
[14]Cohn, A. E.: Jour. Exper. Med., 1912, xvi, 732; Robinson, G. Canby: Ibid., 1913, xvii, 429; Cohn and Lewis, T.: Ibid., 1913, xviii, 739.
[15]Mathison, G. C.: Jour. Physiol., 1910, xli, 416.
[16]Porter, W. T.: Am. Jour. Physiol., 1911, xxvii, 276; ibid., 1915, xxxvi, 418.
[17]Martin, E. G., and co-workers: Am. Jour. Physiol., 1914, xxxii, 212; xxxiv, 220; 1915, xxxviii, 98; 1916, xl, 195.
[18]Bayliss, W. M.: Proc. Roy. Soc., 1908, lxxx, B, 339.
[19]Hill, Leonard: The Physiology and Pathology of the Cerebral Circulation, J. and A. Churchill, 1896.
[20]Hill, L., and Macleod, J. J. R.: Jour. Physiol., 1900, xxvi, 394.
[21]Macleod, and Pearce, R. G.: Am. Jour. Physiol., 1914, xxxv, 87.
[22]Porter, W. T.: Am. Jour. Physiol., 1898, i, 144.
[23]Hill, L., and Barnard, H.: Jour. Physiol., 1887, xxi, 323.
[24]Carter, E. P.: Jour. Lab. and Clin. Med., 1916, i, 719.
[25]Elliot-Smith, G., and Pear, T. H.: Shell Shock, Longmans, Green & Co., 1917.
[26]Porter, W. T.: Am. Jour. Physiol., 1907, xx, 399.
[27]Seelig, M. G., and Joseph, D. R.: Jour. Lab. and Clin. Med., 1916, i, 283; also Seelig and Lyon, E. P.: Surg., Gynec., and Obst., 1910, ii, 146.
[28]Henderson, Yandell: Am. Jour. Physiol., 1908, xxi, 155; also Mann: Bull. Johns Hopkins Hosp., 1914, p. 210; Markwald, J., and Starling, E. P.: Jour. Physiol., 1913, xlvii, 275.
[29]Morrison, R. A., and Hooker, D. R.: Am. Jour. Physiol., 1915, xxxvii, 86.
[30]Pike, F. H., Stewart, G. N., and Guthrie, C. C.: Jour. Exper. Med., 1908, x, 499; see also Dolley, D. H.: Jour. Med. Research, 1909, p. 95, and 1910, p. 331.
[31]Janeway, H. H., and Jackson, H. C.: Proc. Soc. Exper. Biol. and Med., 1915, xii, 193; Erlanger, J.: Gesell, Gasser, Proc. Am. Physiol. Soc., Am. Jour. Physiol., 1918, xlv.

[32]Henderson, Y., and Haggard, W. H.: Jour. Biol. Chem., 1918, xxxiii, 333, 345-355-365 (gives older references). See also Macleod, J. J. R.: Jour. Lab. and Clin. Med., (editorial), 1918, iii.
[33]Short, Rendel: Lancet, London, 1914, p. 131.
[34]Mann: Jour. Am. Med. Assn., 1918, lxx, 611. Also Boston Med. and Surg. Jour., 1917.
[35]Cannon, W. B.: Papers by Cannon and Collaborators in Jour. Am. Med. Assn., 1918, lxx, 520, 526, 531, 611, 618.
[36]Porter, E. L.: Proc. Am. Physiol. Soc., Am. Jour. Physiol., 1916, xlii, 606.
[37]Wiggers, C. J., and Dean, A. L.: Am. Jour. Physiol., 1916, xlii, 476; Am. Jour. Med. Sc., 1917, clii, 666.

PART IV

THE RESPIRATION

CHAPTER XXXIV

RESPIRATION

For convenience, the physiology of respiration may be considered under its *mechanics*, its *control*, and its *chemistry*.

THE MECHANICS OF RESPIRATION

Of the many factors concerned in maintaining the normal functioning of the animal body, the respiratory act is probably the most important. On this account and also because we are conscious of the respiratory movements, the physiology of respiration has been studied from the earliest times. Much of the earlier work naturally concerned itself with the study of the air that enters and leaves the lungs at each respiration—the ventilation of the lungs, as it may be called. Two obvious properties of the respired air are: (1) its pressure and (2) its volume.

The Pressure of the Air in the Respiratory Passages—the Pulmonary or Intrapulmonic Pressure

This is readily measured by inserting a tube into one nostril and connecting the tube with a manometer; at each normal inspiration the manometer registers a negative pressure of 2 or 3 mm. Hg, and at each expiration, a positive pressure of about the same degree. Although normally of small magnitude, the intrapulmonic pressure may become very great when any obstruction is offered to the free passage of the air. The greatest possible expiratory pressure can be measured by simply blowing into a mercury manometer, when it will be equal to that which all the muscles of the thorax and abdomen can exert in compressing the lungs. In a strong man it may amount to more than 100 mm. Hg. Similarly, the greatest possible negative pressure on inspiration may be measured by attempting to inspire through a tube connected with a manometer. It represents the force with which the musculature

of the thorax and abdomen can open up the thoracic cage, and may equal –70 mm. Hg. These measurements in themselves are not of much importance, except as a measure of muscular development.

Intrapulmonic pressures that are intermediate between the two extremes will be acquired in the lower air passages in cases in which there is partial obstruction of the upper respiratory passages, as in bronchitis, spasm of the glottis, diphtheria, etc. During *coughing* also, the intrapulmonic pressure may become very high. In this act the thorax is first filled with air by a deep inspiration; the glottis is then closed, and a forced expiration is made. When a sufficiently high intrapulmonic pressure is attained, the glottis opens and the sudden change in pressure causes so forcible a blast of air that the offending foreign substance is frequently carried with it out of the air passages. It is often assumed that during coughing the sudden increase in pressure in the alveoli will tend to cause their walls to rupture. This, however, is not the case. The alveoli do not alone support the increase of pressure; they merely act as the inner layer of a practically homogeneous structure composed of lung, pleura and thoracic cage. When the tissues of the lung are partially degenerated or atrophied, as in old people, then it is possible that a rupture may take place, but under ordinary conditions it is not likely to occur.

Amount of Air in the Lungs

Measurements of the amount of respired air have recently assumed a considerable interest on account of the various applications which can be made of them in the study of lung conditions. The *tidal air* is that which enters and leaves the lungs with each respiration (about 500 c.c.); the *complemental air* is that which we can take in over and above an ordinary tidal respiration (about 1500 c.c.); and the *supplemental air,* is that which we can give out after an ordinary tidal expiration (about 1500 c.c.). Taking these three together, we have what is known as the *vital capacity.* It is usually about 3500 c.c., and is represented by the amount of air which we can expel from the lungs after as deep an inspiration as possible. The vital capacity is diminished in certain pulmonary diseases (see page 314). After all the supplemental air has been expelled, there still remains in the lungs a large volume of air which can not be voluntarily expelled. This is known as the *residual air.* To measure it in a dead animal it is necessary to clamp the trachea, open the thorax, remove the lungs to a vessel of water, and then allow the air to collect from the opened trachea in an inverted graduated cylinder. One part of the residual air is sometimes called the *minimal air;* it is

represented by that which is not expelled from the lungs of a dead animal when the thorax is opened. In the collapse of the lungs thus produced, the alveoli are not completely emptied of air, because some becomes pocketed within them and is expelled only when the lungs are compressed under water.

The volume of the residual air can readily be measured *during life* by causing a person, after a forced expiration, to take two or three breaths in and out of a rubber bag containing a measured quantity of an indifferent gas such as hydrogen. Suppose the bag to contain at the start 4000 c.c. of hydrogen, and after a few breaths 3000 c.c. of this gas and 1000 c.c. of other gases (the total volume of hydrogen and expired air in the bag being still 4000 c.c.); then the residual air will

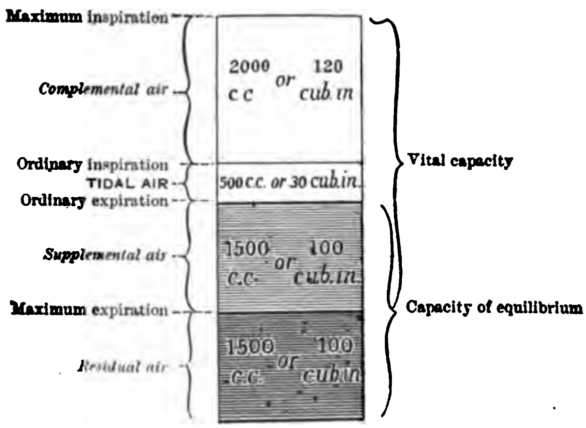


Fig. 107.—Amounts of air contained by the lungs in various phases of ordinary and of forced respiration. (From Waller.)

be 1333 c.c., for it is evident that after a few breaths the composition of the expired air in the bag will be the same as that in the lungs. This calculation is based upon the assumption that no hydrogen is absorbed by the blood during the experiment, which is not strictly the case. The amount absorbed is, however, so small in two or three breaths as to make it permissible to disregard it. The measurement can also be made by taking a few breaths in and out of a bag containing pure O_2. By ascertaining the proportion of nitrogen that collects in the bag, the quantity of residual air can be calculated. We shall see later that the measurement of the residual air during life has some practical importance in connection with the measurement of the bloodflow through the lungs.

Alveolar and Dead Space Air

In addition to these moieties of respired air, we have to consider the division of the air in the lungs into what is called *alveolar* air and *dead-space* air. The former is the air which comes in contact with the epithelium through which gas diffusion between the blood and the air occurs, the latter being the air which fills the respiratory passages. The dead space can not be defined anatomically with exactitude; it is functional rather than morphologic.

Measurement of the volume of the alveolar and dead-space air can be made in an animal breathing under normal conditions by taking advantage of the fact that, while it is in the lungs, the air has added to it CO_2 gas, which is present in the inspired air only in negligible traces. The necessary data are: (1) the volume of the tidal respiration; (2) the percentage of CO_2 in alveolar air; (3) the percentage of CO_2 in the tidal air. Suppose the values to be 500 c.c., 6 per cent and 4 per cent, respectively; then the volume of alveolar air must be $500 \times \frac{4}{6} = 333$ c.c., and the dead space 167 c.c. The measurement so made is accurate only when certain precautions are taken. Because of the practical importance of this part of our subject we shall, however, defer its further consideration until we have become familiar with the general features of pulmonary physiology. Since the first air to move into the alveoli at the beginning of inspiration is that present in the dead space,—the last air expelled from the alveoli on the previous expiration,—it is of no value in purifying the air already present in the alveoli. If we take a tidal inspiration as amounting to 500 c.c. and the functional dead space as 150 c.c., it is plain that only 350 c.c. of the outside air gains the alveoli, and that the subsequent expiration is composed of 150 c.c. of outside air that had lodged in the dead space plus 350 c.c. of alveolar air.

These facts deserve a certain amount of emphasis because of their practical importance in many phenomena connected with respiration. One seldom thinks, for example, that out of the 500 c.c. of air inspired with each breath, only 350 c.c. reaches the alveoli, where it comes in contact with the 2500-3000 c.c. of air already present in this part of the lungs.

There must therefore be a sort of *interface* somewhere in the alveoli between the fresh outside air that comes in with each breath through the bronchioles and the air which is more or less stagnant in the alveoli. This interface must move backward and forward somewhat with each breath, and a rapid diffusion of oxygen and of CO_2 must take place

across it between the inspired air and that in the alveoli. It is impossible to fix any anatomic point at which the interface occurs.

The above described mechanism for the ventilation of the alveoli insures the maintenance of slight but constant changes in the composition of the air next the alveolar epithelium. It helps to prevent sudden variations in the amount of gases in the blood, particularly of CO_2. Should such variations occur, irregular stimulation of the respiratory and other important centers that are influenced by the amount of this gas present in simple solution in the blood, would be the result. *The mechanism serves as a sort of mechanical buffer by diminishing the sudden changes in gas concentration produced by inspiration and expiration.*

Respiratory Tracings

The measurements of air for the determination of the foregoing values are made by the use of meters of various types. Sometimes, however, it is necessary to obtain an inscribed record of the respirations.

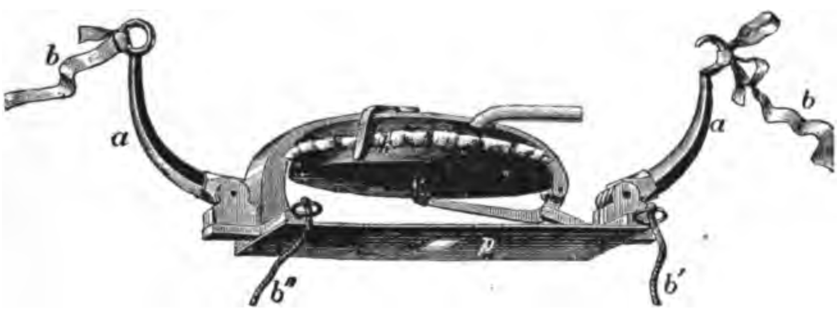


Fig. 108.—Pneumograph. The straps (b, b) are held around the thorax, and the tube of the tambour connected by rubber tubing with a recording tambour.

This may be either qualitative or quantitative. A *qualitative record* is taken by attaching some sort of receiving tambour to the thoracic wall (the best type is shown in Fig. 108), and connecting this with a recording tambour arranged to write on a blackened surface. When it is desired merely to count the respirations or to observe their regularity, such a tracing is all that is required, but obviously it does not tell us *how much air* has entered and left the lungs at each respiration. To obtain a *quantitative tracing,* we must either connect a recording instrument with the trachea or inclose the body of the animal in what is known as a body plethysmograph. In observations on laboratory animals the best type of recording instrument to connect with the respiratory passages is the Gad or Krogh pneumograph. A body plethysmograph as used in the case of man is shown in Fig. 109. All these instruments must of course be calibrated, which is done by pouring a definite num-

ber of c.c. of water from a graduate into a bottle with which the recording instrument is connected by tubing. The displacement of the writing point gives us the necessary data for standardization.

The Intrapleural Pressure

The air which we have just been considering depends for its movement in and out of the air passages upon changes occurring on the outer aspect of the lungs in the space between them and the thoracic wall. This is called the intrapleural space. It does not really exist as an actual space in the living animal, for the visceral pleura which covers the lungs is in accurate and intimate apposition with the parietal pleura on the inner aspect of the thorax.

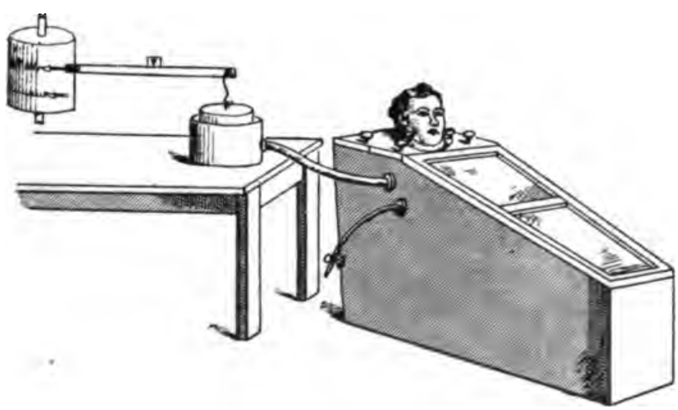


Fig. 109.—Body plethysmograph for recording respiration. (From J. S. Haldane and J. G. Priestley.)

If the thoracic walls are punctured in a living animal or in one which has recently died, the air will rush into the thorax, the two layers of pleura separate, and the lungs collapse, causing temporarily a space to be formed between the two layers of pleura and indicating that a certain subatmospheric or negative pressure must exist in the intact thorax to prevent the lungs from collapsing. The degree of this negative pressure may be measured by connecting a tube and a manometer with the thoracic cavity. While the thorax is at rest, as in expiration or immediately after death, this pressure amounts to about –5 millimeters.* On inspiration it increases to –10 millimeters. There are therefore two problems to be considered: (1) the cause of the negative pressure in the quiescent thorax, and (2) the cause of the increase of the negative pressure during inspiration.

*The minus sign indicates that the pressure is negative or subatmospheric. It is a suction pressure.

The Permanent Negative Pressure.—Let us start with the changes that occur in the thorax when the first breath is drawn. While the animal is still *in utero*, the lungs completely fill the thorax. When the first breath is drawn the thoracic cage expands more quickly than the lungs, so that the latter become stretched, the stretching force being the air that is introduced into them from the outside through the trachea and bronchial tubes. On becoming stretched the lungs fill the increased space created in the thorax by the greater expansion of the thoracic cage. This in itself, however, would not explain the cause of a subatmospheric pressure in the intrapleural space. Another factor must come into play—namely, the elastic tissue of the lungs, which by the expansion will become stretched and, therefore, tend constantly to relax to its previous condition and so exert a pull on the structures between it and the thoracic wall. It is this elastic recoil which we really measure when we connect a manometer with the intrapleural space. Throughout life the lungs remain of smaller size than the thoracic wall, and therefore to fill the thoracic cavity they are constantly more or less distended and the elastic tissue somewhat stretched. The lungs are, however, not the only structures in the thorax which become expanded; all thin-walled vessels and viscera, like the veins, the esophagus, the auricles, etc., must also become opened out a little.

When the thoracic wall is punctured and the outside air allowed free entry to the intrapleural space, differences in pressure no longer exist on the inner and outer aspects of the lungs, so that they collapse into the postmortem condition on account of the elastic recoil. If a puncture in the thoracic wall of a living animal is immediately occluded, the lungs will expand again, because the blood absorbs the gases from the intrapleural space and recreates the partial vacuum required to expand the lungs. This absorption of gas in the pleural cavity is usually quite rapid; but if the *pneumothorax*, as the condition is called, is allowed to persist for any length of time, the lungs will not become properly expanded again.

The Greater Negative Pressure on Inspiration.—The cavity of the thorax becomes increased in all diameters during inspiration, with the result that a greater space in the pleural cavity has to be filled. All the thin-walled structures in the thorax therefore become still more stretched, the lungs of course participating to the greatest extent because of the entrance of outside air. The stretching of the elastic structures causes a greater pull, or negative pressure, to be exerted in the pleural cavity. Instead of being -5 mm. Hg, as in expiration, the intrathoracic pressure now comes to be above -10 mm. Hg.

When any *obstruction* exists in the air passages, the changes in intra-

thoracic pressure produced by the movements of respiration become more pronounced than under normal conditions. When the thorax expands with the trachea blocked, the lungs are not able to open up sufficiently to fill all the space so that there is excessive dilatation of the veins, auricles and esophagus, as well as drawing in of the intercostal spaces and bulging upwards of the diaphragm. If a manometer is connected with the pleural space under these conditions, a very large negative or suction pressure will be observed, amounting often to –70 or –80 mm. Hg. It is possible that under such conditions some space might temporarily exist between the parietal and visceral layers of the pleura, but it could not remain long, for it would very soon be filled by fluid exuding from the blood vessels. In the opposite condition, in which the respiratory passages are blocked and a forced expiration is made, as for example in the first stage of coughing or during such acts as defecation and parturition, the thoracic cage is compressed upon the viscera, with the result that the air in the lungs assumes a positive pressure, amounting often to nearly 100 mm. Hg. If a puncture wound is made in the thorax under these conditions, the lungs instead of collapsing will bulge out of the wound, for what is really occurring is that the thorax is forcibly contracting on occluded sacs of air.

It is the alternating changes in intrapleural pressure that are responsible for the changes in intrapulmonic pressure and these for the movement of air in and out of the lungs with each respiration. In other words, the thorax does not expand on inspiration because air rushes in, as the uninitiated imagine, but air rushes in because the thorax expands.

The Influence of Intrapleural Pressure on the Blood Pressure.—The movements of respiration produce effects on the vascular system that are of considerable importance in maintaining the circulation of the blood. If an arterial blood-pressure tracing is examined, it will be observed that aside from the cardiac pulsations large waves exist on it that are approximately synchronous with the respiratory movements, the upstroke of each of these waves corresponding in general with inspiration, and the downstroke with expiration (Fig. 22). These respiratory variations in blood pressure might be due either to changes in heart rhythm or to a purely mechanical cause. Regarding the first possibility, it is indeed the case in most animals that the pulse is quicker on inspiration than on expiration, but that this alone is not an adequate explanation of the rise is shown by the fact that it still persists after the vagus control of the heart has been eliminated, either by cutting the nerve or by the action of atropine.

The cause must therefore be a mechanical one. Bearing in mind the

effects which we have seen are produced on the movement of air in and out of the lungs by the changes in capacity of the thorax with each respiration, we naturally assume that the increase in blood pressure may be due to the fact that on inspiration more blood is sucked out of the systemic veins into those of the thorax, that this excess when it is propelled by the heart into the arteries raises the blood pressure, and that on expiration the opposite condition obtains. That the movements of the thorax on inspiration do accelerate the speed with which the venous blood is traveling towards the heart can easily be shown by measurements of bloodflow.

This explanation, however, does not suffice to account for all the changes of blood pressure which occur in respiration, for if we take very accurate tracings of blood pressure and of the respiratory movements side by side, we shall find that, although, in general, the blood pressure rises with inspiration, yet the beginning of the rise is considerably delayed; that is, immediately following the beginning of the inspiratory act the arterial blood pressure continues for some time to fall, and at the beginning of expiration it continues for some time to rise (Fig. 22). Moreover, it will be found, if tracings taken from different animals are compared, that frequently the general effect of expiration is to cause more rise than fall, and of inspiration more fall than rise. It will be found that these differences are dependent largely on the type of respiration, whether thoracic or abdominal (Lewis).[11]

Let us consider first of all exactly what will happen in an animal breathing entirely by the *thorax* (e.g., the rabbit). The first effect of the inspiration is to cause the veins leading to the auricles, the auricles themselves and the blood vessels of the lungs to become suddenly expanded. More blood therefore will flow into them. For a moment or two this blood will, however, tend to stagnate in the more capacious vessels, and it will be some time until it finds its way to the left side of the heart; therefore the initial effect of inspiration is a distinct fall in arterial blood pressure. When the extra space created in the blood vessels has been filled with blood,—that is, when inspiration has practically ceased,—the blood will flow on in increased volume to the left side of the heart, and, therefore, raise the arterial blood pressure. On *expiration* the first effect is that the diminishing negative pressure will cause the thin-walled vessels mentioned above to constrict and thus squeeze the blood inside them into the left side of the heart and raise the pressure; but the ultimate effect in the later stages of expiration will be that the vessels, being constricted, will allow less blood through them and the arterial blood pressure will fall.

Take now the case of *abdominal respiration*. In inspiration the dia-

phragm descends and crowds the viscera against the vena cava, with the result that at first more blood is squeezed into the thorax and the blood pressure tends slightly to rise. After this initial effect, however, the compression of the vena cava causes less blood to reach the thorax, and the arterial blood pressure falls. The conditions will be exactly reversed on expiration. The initial effect of thoracic inspiration is, therefore, to make the arterial blood pressure fall, and the initial effect of abdominal inspiration, to make it rise. The net effect

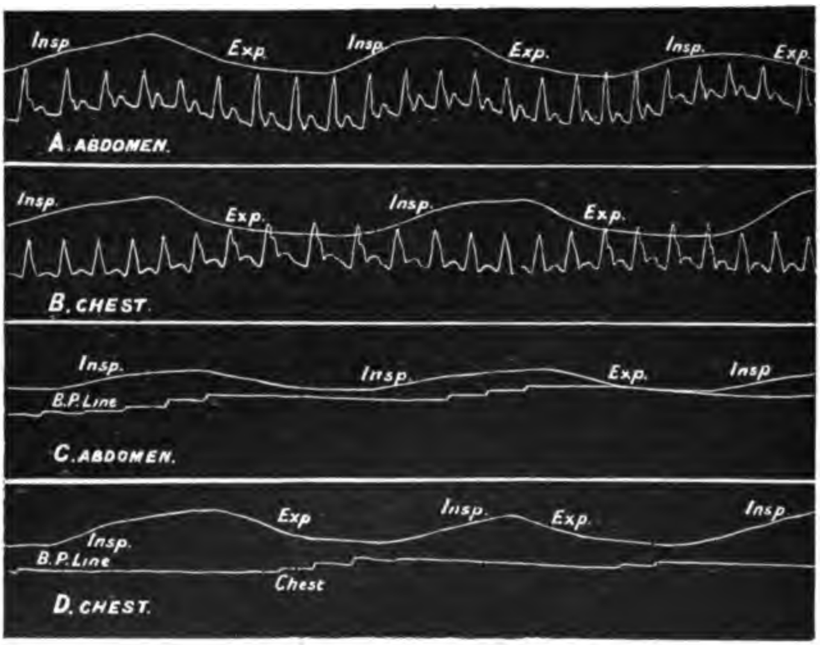


Fig. 110.—Effect of abdominal and chest breathing on the pulse and blood pressure of man. Abdominal inspiration raises the pressure and diminishes the amplitude of the pulse curve. Thoracic inspiration less clearly lowers the pressure. Expiration has the opposite effects. (From Lewis.)

produced will be the algebraic sum of these two opposing influences (see Fig. 110).

Another factor that comes into play in determining the effect of the respiratory movements on the cardiac output acts through the changes in the pericardial pressure. When this is lowered, as early in inspiration, it encourages diastole, thus causing better filling and therefore better discharge from the heart.

These considerations taken together make it easy to understand the changes in blood pressure, particularly in the veins, which occur when a forced inspiratory or expiratory movement is made with the glottis closed. *A forced expiration* of this nature occurs during the acts of

defecation and parturition, as well as in the first stages of coughing; it is also produced by blowing into a tube, or against some resistance. On account of the positive pressure that is brought to bear on the veins as they enter the thorax, the venous pressure suddenly rises, slowing down the flow of blood through the capillaries and causing bulging of the veins and, if the effect is sustained, cyanosis. On the arterial side of the vascular system, after a momentary rise caused by the squeezing out into the left side of the heart of the blood in the capillaries of the lungs, there is a more permanent fall in pressure due to the fact that less blood is now getting from the right side to the left side of the heart. After some time the pressure begins to rise again, partly on account of the back pressure through the capillary vessels and partly because of vasoconstriction as a result of asphyxial conditions.

In the opposite condition, during *a forced inspiratory movement* with the glottis closed or with the mouth attached to some tube through which the attempt is made to suck air, the thoracic cavities open up without the lungs being able to occupy completely the extra space. The dilatation of the veins and other thin-walled structures in the thorax thus causes an immediate fall in both the venous and the arterial pressure—in the venous, because the blood is sucked toward the large vessels in the thorax and lungs, and in the arterial, because the blood is now delayed in its passage from the right to the left side of the heart. If this condition is maintained, the arterial pressure may recover somewhat, but that in the veins is permanently lowered.

CHAPTER XXXV

THE MECHANICS OF RESPIRATION (Cont'd)

VARIATIONS IN THE DEAD SPACE, THE RESIDUAL AIR AND MID-CAPACITY, AND THE VITAL CAPACITY IN VARIOUS PHYSIOLOGIC AND PATHOLOGIC CONDITIONS

BY R. G. PEARCE, B.A., M.D.

Dead Space

Under ordinary conditions of breathing the dead space is fairly constant in volume. Haldane[5] and Henderson[6] believe that it may be increased by 400 per cent in maximal deep breathing, and that the increase is due to the passive stretching of the lower air sacs. Although such large variations in the capacity of the dead space has not been observed by Krogh and Lindhard[7] or by R. G. Pearce,[8] it is undoubted that moderate rhythmic variations may occur. Even in deeper breathing (1500 c.c. or over), a slight increase, which with maximum breaths may amount to 100 c.c., can be demonstrated. This is not surprising when we remember that the walls of the bronchi and bronchioles are made up largely of readily expansible tissue (elastic and smooth-muscle fibers). As the respirations become deeper and the expanding force of the inspiratory movements of the thorax becomes more pronounced, the diameter of the bronchi and bronchioles will enlarge proportionately—that is, the diameter or circumference will increase in direct proportion to this force; but the area of the cross section of the bronchi (i. e., the capacity) will increase as the square of the diameter. This depends on the fact that the area of a circle is increased by 125 per cent when the diameter is increased by 50 per cent, and by about 300 per cent when the diameter is increased by 100 per cent.

The capacity of the dead space has a certain clinical significance. Siebeck[9] has estimated that the dead space may increase by 100 c.c. in asthma, but others believe that the increase may be greater. One reason for the discordant results lies in the fact that the percentage of CO_2 found in the alveolar air obtained by the Haldane-Priestley method has been used as one of the basic figures in the determination of the

capacity of the air passages. As explained elsewhere (page 344), the prolongation of expiration required to obtain the sample of alveolar air by this method gives figures that are too high even under normal conditions, and it is plain that this error will be exaggerated in asthma, where the expiration is greatly prolonged. An increase in the capacity of the dead space must be accompanied by an increase in the respiratory volume if the alveoli are to be adequately ventilated. It has been thought by some clinicians that the difficulty in asthma, emphysema and cardiac decompensation may lie in part in an increase in the dead space. Careful estimations of the dead space in these conditions, however, fail to demonstrate any great variation.

An explanation of the fact that the dead space in emphysematous patients has been found to be generally large when determined by the Haldane-Priestley method (see page 340), and also for some of the clinical phenomena accompanying the condition, may be as follows: In emphysema the walls of the alveoli, especially about the lateral and lower borders of the lungs, have lost their elasticity and fail to expand or relax properly during the respiratory cycle. As a result the air in these alveoli remains relatively unchanged except when forced respirations are made. When a sample of alveolar air is taken directly, this dead air is pushed out of the distended and diseased alveoli by the forced respiration required in the direct sampling of the alveolar air. Since the air in these alveoli has been in contact with the blood entering the lungs, it has a high CO_2 content, which results, when compared with the uniformly low CO_2 content found in the tidal air, in giving a large figure for the dead space. Since the capacity of the dead space is not increased, the blood in the normal alveoli is probably being superventilated in order to compensate for the high CO_2 tension in the blood entering the left heart from the diseased alveoli. However, the O_2 content of the blood leaving the sound alveoli is practically normal (because superventilation can not cause it to take up more), and can not compensate for the low O_2 content in the blood coming from the diseased alveoli, the net effect being therefore a low tension of O_2 in the blood leaving the heart, which accounts for the cyanosis often seen in emphysema (Pearce). A somewhat similar explanation can be given for the cyanosis present in pulmonary edema, if we assume that all the alveoli in this condition do not share alike in the edema (Hoover).

The Residual Air and Mid-capacity of the Lungs

During muscular exercise the residual air of the lungs is increased, and the vital capacity decreased (Bohr). This causes the lungs to as-

sume a more inflated condition between breaths or, as it has been clumsily styled, a greater mid-capacity. These changes may serve as a physiologic method for increasing the efficiency of alveolar ventilation so as to meet the greater needs of the body. This is partly because the pulmonary vessels become dilated and the bloodflow through the lungs is favored, and partly because of the influence of the reserve and supplemental airs on the tension of the arterial blood gases during the respiratory cycle. For example, if the lungs were completely depleted of air during expiration, the blood leaving them at the end of this act would be entirely venous. On the other hand, if the amount of air left in the lungs at the end of expiration were above the normal amount, each increment of CO_2 given off from the blood, or of O_2 absorbed by it would produce less change in the pressure of the CO_2 or O_2.

The importance of these influences will be seen from the following figures. If the residual and supplemental air amounts to 2000 c.c., and the percentage of CO_2 in the alveolar air at the end of expiration is 5 per cent, then 100 c.c. of CO_2 must be present in the lungs. In a condition of bodily rest about 20 c.c. of this gas is excreted during a respiratory cycle, so that if the breath were held during this period, the percentage of CO_2 would rise from 5 to 6 per cent, and an inspiration of 400 c.c. would be required to bring the air in the lungs back to 5 per cent of CO_2. On the other hand, if the residual and supplemental air amounted to 3000 c.c. with 5 per cent of CO_2 in the alveolar air at the end of the expiration, there would be 150 c.c. of CO_2 in the lungs at the end of the expiration, so that holding the breath for the time of the respiratory cycle would raise the percentage of CO_2 only to 5.66 (provided the production of CO_2 was the same as in the first case), and an inspiration of 600 c.c. would be necessary to reduce it to the normal expiratory figure. Or, putting it another way, the production of CO_2 can be increased 50 per cent in the time of a respiratory cycle without affecting the tension of gases in the lungs, provided the residual and supplemental air and the volume of the respiration are increased 50 per cent. If only one of the factors is changed, however, then the balance of the respiration must be disturbed, and the greater variation in the tension of the gases in the arterial blood must occur at the different phases of the respiratory cycle. Bohr and Siebeck have shown that the residual air is invariably increased in emphysema and that the mid-capacity of the lungs is likewise increased; and it would appear from Siebeck's data that a similar condition must be present in cases of decompensated heart.

Patients suffering from dyspnea, particularly those suffering from

cardiac dyspnea, can not breathe as comfortably when lying as when sitting. This condition is known as *orthopnea.* The advantage of the sitting over the lying position for breathing can not be satisfactorily explained. The greater vital capacity in the upright position; the favoring of the return of the venous blood from the cerebral vessels by gravity; the increased caliber of the pulmonary vessels because of the enlarged thoracic cavity (see page 318); and the increase in the reserve air of the lungs—are all factors to be considered.

The Vital Capacity.—At one time it was thought that the vital capacity of the lungs was related to their ventilatory capabilities, but for years the determination of this value in patients has been considered unimportant. Recently Peabody and Wentworth[10] have called attention to the fact that patients with heart disease become dyspneic more readily than do healthy subjects, and that this tendency seems to depend largely on their inability to increase the depth of the respiration in a normal manner. They find that this inability to breathe deeply corresponds to a diminished vital capacity of the lungs as measured in a spirometer, by the volume of the greatest possible expiration after the deepest inspiration. They believe that any condition which limits the possibility of increasing the minute volume of air breathed must be an important factor in the production of dyspnea.

In normal adults the following averages (Table I), were secured from a large series of clinical cases. The subjects are grouped into two classes, each group being subdivided according to height.

TABLE I

THE VITAL CAPACITY OF THE LUNGS OF NORMAL MALES

GROUP	NUMBER STUDIED	HEIGHT IN FEET AND INCHES	NORMAL VITAL CAPACITY C.C.	NUMBER WITHIN 10% OF NORMAL	HIGHEST VITAL CAPACITY	LOWEST VITAL CAPACITY	HIGHEST %	LOWEST %	NUMBER BELOW 90% OF NORMAL
I	14	6′ +	5,100	9	7,180	5,030	141	99	0
II	44	Over 5′ 8½″ to 6′	4,800	41	5,800	4,300	121	90	0
III	38	5′ 3″ to 5′ 8½″	4,000	31	5,080	3,450	127	86	1
THE VITAL CAPACITY OF THE LUNGS OF NORMAL FEMALES									
I	10	Over 5′ 6″	3,275	5	4,075	2,800	124	86	2
II	13	Over 5′ 4″ to 5′ 6″	3,050	9	3,425	2,660	112	88	2
III	21	5′ 4″ or less	2,825	16	3,820	2,500	135	89	1

(Peabody and Wentworth.)

It would appear that in normal people the vital capacity is at least 85 per cent, and almost always 90 per cent or more, of the standard adopted for each group. In elderly persons a slight decrease from these standards may be expected.

TABLE II

THE RELATION OF THE VITAL CAPACITY OF THE LUNGS TO THE CLINICAL CONDITION IN PATIENTS WITH HEART DISEASE*

GROUP	VITAL CAPACITY %	NUMBER OF CASES	MORTALITY %	SYMPTOMS OF DECOMPENSATION %	WORKING %	REMARKS
I	90 –	25	0	0	92	Few symptoms referable to heart.
II	70 to 90	41	5	2	54	History of dyspnea with exertion, yet able to do moderate work.
III	40 to 70	67	17	89	7	Dyspnea with moderate exercise. Few able to work.
IV	Under 40	23	61	100	0	Bedridden, with marked signs of cardiac insufficiency.

(Peabody and Wentworth.)

*Certain cases were tested several times and, owing to changes in the vital capacity they appear in more than one group. In the "mortality" column they are included only in the lowest group into which they fell. "Symptoms of decompensation" indicate dyspnea while at rest in bed or on very slight exertion. Under "working" are included only those actually at work, and able to continue. Many other patients in Group II were able to work, but they are not included as they were still in the hospital.

Table II shows that there is a remarkably close relationship between the clinical condition of cardiac patients, particularly as regards the tendency to dyspnea, and the vital capacity of the lungs. Peabody and Wentworth believe that the determination of the vital capacity affords a clinical test as to the functional condition of the heart, since compensated patients who do not complain of dyspnea on exertion have a normal vital capacity. Patients with more serious disease in whom dyspnea is a prominent symptom, have a low vital capacity; and the decrease in vital capacity runs parallel with the clinical condition. As a patient improves, his vital capacity tends to rise; as he becomes worse, it tends to fall. In other diseases in which mechanical conditions interfere with the movements of the lungs, the tendency to dyspnea corresponds closely to the decrease in the vital capacity. The cause of the decrease in the vital capacity of the lung in cardiac decompensation is difficult to explain satisfactorily. It may be the limitation in the movements of the lungs produced by engorgement of the pulmonary vessels, by the weakness of the intercostal muscles, the rigidity of the bony thorax, emphysema, or accumulation of fluid in the pleural cavities.

In cardiac disease the air in the lungs at the end of a normal expiration is usually increased. This is similar to the condition which attends exercise, and is probably a physiologic adaptation to give optimum aeration to the blood, as explained above.

CHAPTER XXXVI

THE MECHANICS OF RESPIRATION (Cont'd)

THE MECHANISM BY WHICH THE CHANGES IN CAPACITY OF THE THORAX AND LUNGS ARE BROUGHT ABOUT

By R. G. Pearce, B.A., M.D.

The changes that take place in the form and the dimensions of the thorax during respiration are brought about by movements of the ribs, diaphragm, sternum, and vertebræ. The share which each plays must be considered separately.

The Movements of the Ribs

The first seven pairs of ribs progressively increase in length, and are attached directly to the sternum by cartilaginous bands. The eighth to the twelfth pairs progressively decrease in length, and as far as the tenth they are indirectly attached to the sternum by cartilages which join the seventh. The eleventh and twelfth have their anterior ends free, and may be considered a part of the abdominal wall and not an intrinsic part of the thoracic cage.

Each pair of ribs, together with its articulating cartilage and vertebræ. forms a ring, the plane of which is directed forward and downward. The spinal articulations of the upper ribs differ from those of the lower ones. In the former the articulations on the tubercle exist as convex ovoid facets, which fit into corresponding hollow facets on the transverse processes of the vertebræ, while the corresponding facets of the lower ribs are flat. Each transverse process from above downward is tilted a little more backward than the one above, so that the angle at which the ribs are set to the spine increases from above downward. This manner of articulation of the upper ribs with the vertebræ prevents any rotation in the spinosternal axis, so that there can be no so-called bucket-handle movement in this region (Keith). The articulation, however, allows the neck of the rib to rotate in an axis approximately transverse to the body. The angle which the shaft of the rib makes near its neck, together with the arch of the shaft, which is directed downward and forward, has the effect of causing the transverse rotation of the neck of the rib to be

converted into an upward movement, which is greatest in that part of the shaft lying parallel to the axis of rotation of the neck (Fig. 111).

The upper ribs from the first to the fifth form a cone-shaped top to the thorax, whereas the lower ones form a vertical series, each being situated almost directly above its neighbor. The upper set is arranged for the expansion of the conical upper lobe of the lungs, the lower for the expansion of the more or less cylindrical lower lobes. During inspiration the anteroposterior diameter of the conical portion of the thorax increases, because the ribs, together with the sternal connections, move through progressively increasing arches, and each lower rib tends to override the rib just above. The maximal rise of the ribs from the first to the

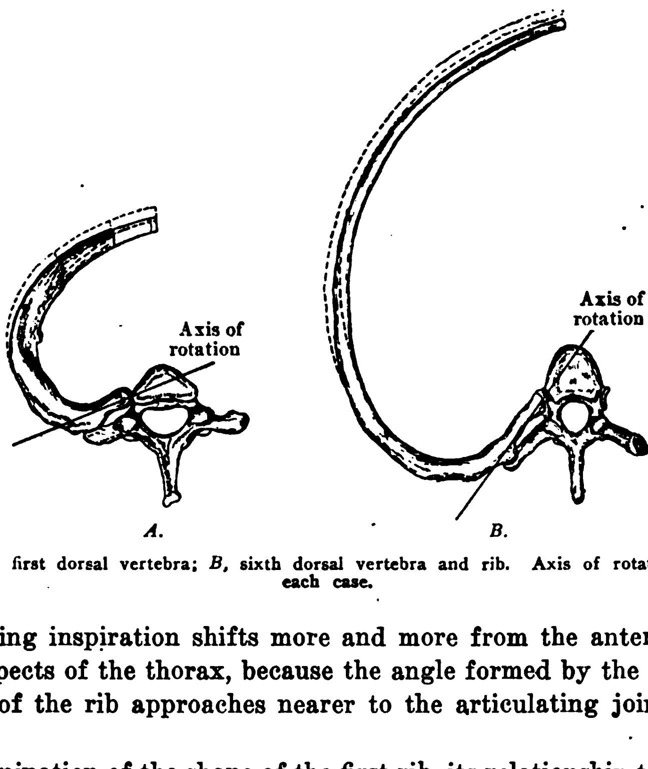


Fig. 111.—*A*, first dorsal vertebra; *B*, sixth dorsal vertebra and rib. Axis of rotation shown in each case.

tenth during inspiration shifts more and more from the anterior to the lateral aspects of the thorax, because the angle formed by the shaft near the neck of the rib approaches nearer to the articulating joints on the vertebræ.

An examination of the shape of the first rib, its relationship to adjacent structures and its movements, shows that it differs from the others in its respiratory function. The first pair of ribs and the manubrium sterni are bound closely together by short, wide costal cartilages, and form a struetural unit which Keith[1] calls the *thoracic operculum*. This lid is articulated behind with the first thoracic vertebra by a joint, which is more nearly transverse than that of the rest of the costal series; and in front with the manubrium, which is also articulated with the clavicles

above and with the body of the sternum below. The freedom of movement at the angle which the manubrium makes with the sternum at this joint is related to the type of breathing. When the lower portion of the sternum is elevated during inspiration, the movement of the joint is not free, but when the sternum is retracted, the movement at the angle may amount to 16°. Lack of movement of the sternal manubrial joint has been considered by some physicians as one of the predisposing causes of pulmonary tuberculosis. During inspiration, the first rib and its anterior attachments are raised by the scaleni, and serve as a point towards which the second, third, fourth and fifth ribs are elevated. During expiration, they are depressed toward the lower ribs, which form a more or less fixed base.

The combined effect of these influences is to produce a motion of the upper ribs which is described by the clinician as being *undulatory*. This movement is more apparent in the upper part of the thorax, because here the relative difference in the length of the ribs is greatest. Hoover attributes a certain diagnostic significance to loss of the undulatory movement, diminution in the extensibility of the underlying lungs causing it to become less or to disappear. The phenomenon is elicited by placing the tip of the ring finger on the second rib in the midclavicular line, the tip of the middle finger on the third rib midway between the midclavicular and anteroaxillary line, and the tip of the index finger on the fourth rib in the midaxillary line. The patient is then instructed to make a moderately rapid and deep inspiration. The finger on the third rib will be observed to move farther than that on the second rib, and the finger on the fourth rib will move farther than that on the third. The movement of each rib from above downward succeeds and exceeds that in the rib just above.

When there is a moderate degree of impairment in the ventilation of the upper lobe, the three ribs move in unison and through the same distance, so that the undulatory movement is lost although the ribs involved may exhibit a considerable excursion. The undulatory movement is also impaired by any disease which encroaches on the air spaces, invades the interstitial tissue of the lung, or displaces the lung as in the case of an enlarged heart or a distended pericardial sac. Another possible factor in this phenomenon is that any inflammatory process in the lung or adjacent tissue will produce a reflex inhibition of the muscles of the ribs, and thus limit the expansion of the thorax.

The axis of movement of the lower ribs, as of the upper ribs, accurately corresponds with that indicated by their articulation with the vertebræ, because the muscles attached to them, as well as the diaphragm, influence their movements to a large extent. Anteriorly the lower ribs from the

sixth to the tenth are joined to the sternum by the cartilages which unite the sixth, seventh, eighth, ninth, and tenth, so that any movement in which the ribs are raised is accompanied by an anterior movement of the sternum (Fig. 112). The ribs are so articulated to the spinal column that the inspiratory act causes the lateral and anterior part of each rib arch to move forward and outward more than the one above it.

In natural breathing in the standing or sitting posture there is a *slight extension of the spine* during inspiration. This serves to increase all diameters of the thorax and its absence is undoubtedly an important

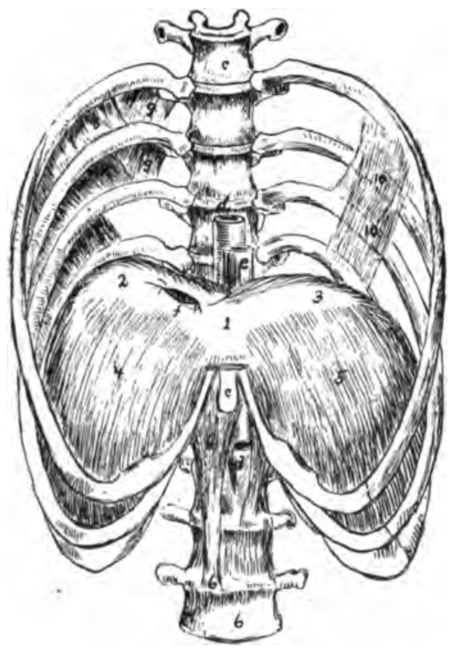


Fig. 112.—Lower half of the thorax from the 6th dorsal to the 4th vertebra, seen from the front. *c*, ensiform process; *d, d'*, aorta; *e*, esophagus; *f*, aperture in tendon of diaphragm for passage of vena cava inferior; *1, 2, 3*, trilobate expansions of tendinous center of diaphragm; *4, 5*, costal portions, right and left, of diaphragm muscle; *6*, right crus of diaphragm; *8, 9*, internal intercostal muscles, which are absent near the vertebral column, where it joins *9* and *9*, the external intercostals; *10, 10*, subcostal muscles of left side. (From Luschka.)

contributory factor in reducing the vital capacity of an individual when lying on the back. Figures given by Hutchinson for the effect which posture has on the vital capacity are of interest because of their bearing on the cause of orthopnea. In the same individual he found the following vital capacities:

Standing	4300 c.c.
Sitting	4200 c.c.
Supine	3800 c.c.
Prone	3620 c.c.

The Action of the Musculature of the Ribs

In a general way, the external intercostal muscles may be considered as a broad extension of the scalene muscles over the thoracic walls, with the ribs as intersections. The scaleni serve to fix the position of the

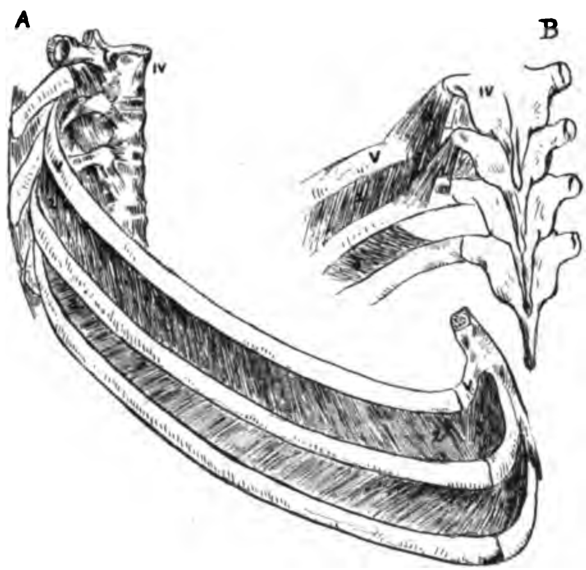


Fig. 113.—Intercostal muscles of 5th and 6th spaces. *A*, side view; *B*, back view; *IV*, 4th dorsal vertebra; *V*, 5th rib and cartilage; *1, 1*, M. levatores costarum, 2, 2, external intercostals; 3, 3, internal intercostals, exposed by removal of the external muscles. In *A*, there are no external intercostals in the intercartilaginous spaces; in *B* there are no intercostals near the vertebral column. (From Allen Thomson.)

first rib so that it forms an anchorage for the action of the external intercostal muscles in raising the lower ribs. They also raise the upper three pairs of ribs along with the manubrium and sternum.

The function of the *intercostal muscles* has been the subject of much

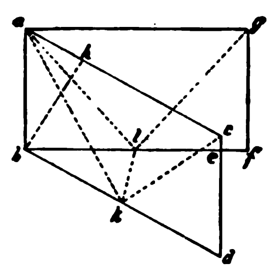


Fig. 114.—Hamberger's schema to demonstrate the functional antagonism of internal and external intercostals.

When the ribs *ac* and *bd* pass into the inspiratory positions *ag* and *bf*, the intercostal space dilates (*bh* is greater than *ab*); the sternum *gf* moves away from the vertebral column *ab* (*bf* is greater than *be*); the fibers of the external intercostals *ak* shorten (*ak* is greater than *al*); and those of the internal intercostals *ck* lengthen (*ck* is greater than *lg*). The reverse occurs when the inspiratory position is taken. (From Luciani's *Human Physiology*.)

debate, and can not be said to be definitely settled. The direction of the fibers in the internal intercostals indicates that they are expiratory in function, since they can not shorten in the inspiratory position; while, on the other hand, the fibers of the external intercostals can not shorten in the expiratory position, and hence must be considered inspiratory in character (Fig. 113). In 1751 Hamberger showed that mechanically this is the case, and gave the schema shown in Fig. 114.

The function of the intercartilaginous muscles, however, must be inspiratory, as is shown in Fig. 115.

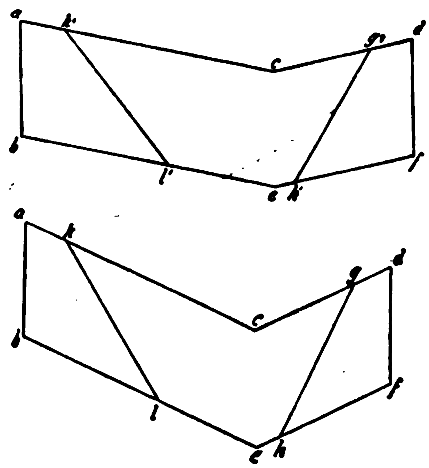


Fig. 115.—Schema to demonstrate that the function of the internal intercartilaginous intercostals is identical with that of the external interosseous intercostals.

The ribs and costal cartilage may be regarded as rods bent at the angles *acd* and *bef*, in which the articular points *c* and *e* represent the symphysis between the bony and the cartilaginous parts on which traction is made. During inspiration the fibers of the intercartilaginous muscles, which have the direction *gh*, move the sternum *df* away from the vertebral column *ab*, like the fibers of the external intercostals, which run in the direction *kl*. During this double action the angles *c* and *e* must be decreased, because the muscles of the upper intercostal spaces work simultaneously, and the entire thorax is slightly elevated during inspiration. From this scheme it is apparent that the external intercostals and the intercartilaginous muscles must be the same. (From Luciani's *Human Physiology*.)

The Action of the Diaphragm

It is possible, however, that the main function of both the intercostal muscles is to regulate the tone of the intercostal spaces and so prevent their suction inwards when the negative pressure in the thorax increases (i. e., suction becomes greater). The ascent of the ribs, while producing an increase in the anteroposterior and transverse diameters of the thorax, would decrease the vertical diameter if this was not counteracted by the fixation of the lower ribs and the descent of the diaphragm. The peripheral edges of the diaphragm are attached behind to the lumbar vertebræ, laterally to the lower edges of the six lower ribs and their cartilages, and in front to the tip of the ensiform cartilage. The fibers converge to

enter the central tendon, and the lateral sheets are pressed upward by the intraabdominal positive and intrathoracic negative pressures, so that they form a dome-shaped vault, with the liver in the right side and the stomach and the spleen in the left.

During expiration the lateral edges of the diaphragm are in contact with the parietal pleura of the thoracic cavity, forming what are known as the pleural sinuses. During inspiration the fibers of the diaphragm shorten; this straightens out the arch of the diaphragm and pulls the lateral edges of the diaphragm away from the parietal pleura, thus opening up the pleural sinuses, into which the lungs descend. Usually the opening up of the sinuses is accompanied by a slight retraction of the external chest wall, which is known as Litten's diaphragm phenomenon. The descent of the diaphragm may produce a movement of from 10 to 15 mm. on each side, which accounts for a rather important fraction of the volume of air exchange by the lungs. The central portion of the diaphragm does not move much in normal respiration, but in forced respiration its movement may be considerable.

Because of its attachments to the lower six ribs, the contraction of the diaphragm tends to pull the margins of the ribs towards the median line, but under normal conditions this movement is opposed by the action of the external intercostals in raising the ribs and expanding the horizontal diameters of the thorax, and by the lower vertebral muscles, which fix the position of the lower ribs.

The relative part which the diaphragm and the external intercostal muscles play in the widening of the lower part of the thorax is of some importance from the standpoint of diagnosis. It has generally been held that the contraction of the diaphragm produces a widening of the lower part of the thorax, because in its descent it presses upon the abdominal viscera and so distends the abdomen and pushes out the lower ribs. That this might occur seems not improbable, but Hoover[2] has recently shown by experimental and clinical observations that the flaring in the costal margins seen in normal inspiration depends on other factors. He calls attention to the fact that the contraction of the intercostals raises the ribs and increases the angular divergence of the subcostal borders. This widening of the angle made by the costal margins at the tip of the sternum is very pronounced in paralysis of the diaphragm while in paralysis of the intercostal muscles, the costal borders are drawn towards the median line and the subcostal angle is decreased. This shows that the diaphragm must tend to diminish the angle.

The line of traction of the diaphragm is a straight one joining the central tendon with the edge of the ribs. When the diaphragm forms a well-defined arch, it exerts its traction at a disadvantage, and the ex-

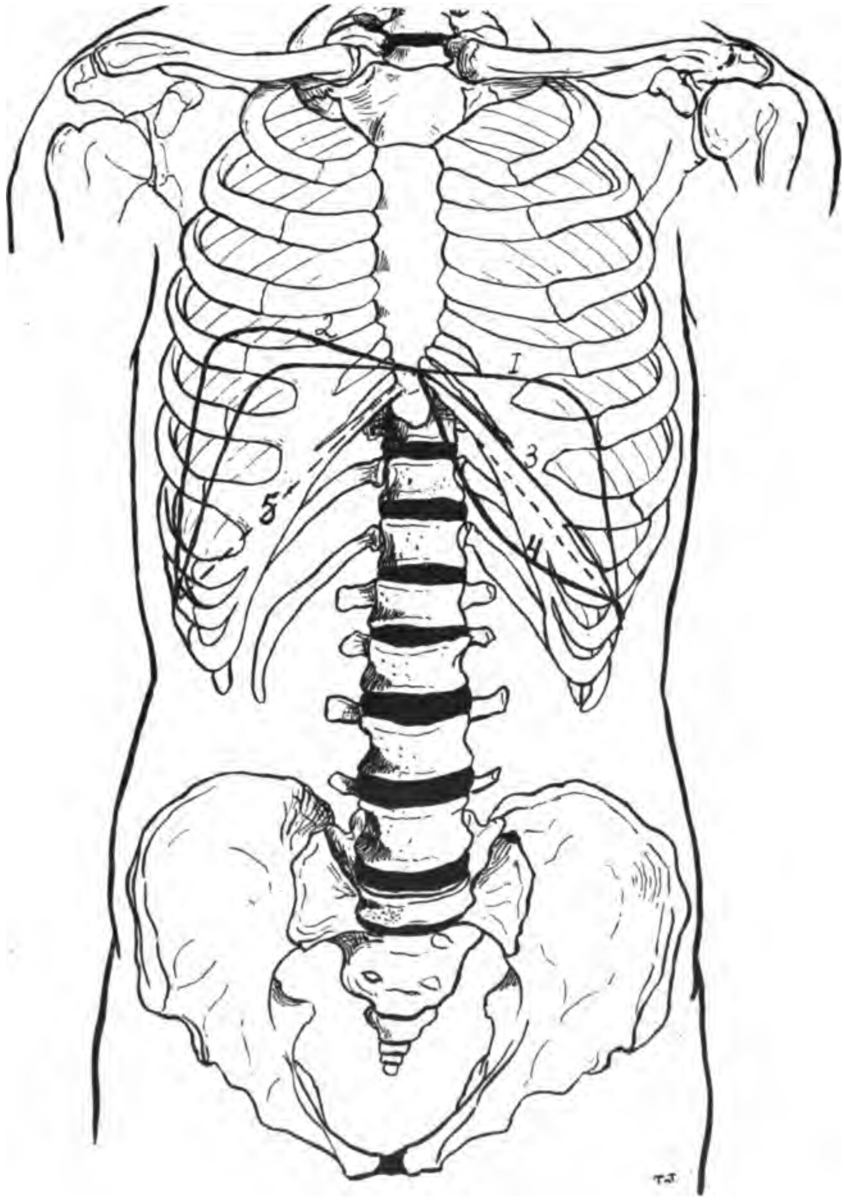


Fig. 116.—Diagram to show the effect of high and low positions of the diaphragm on the costal angle.

Line 1. Normal position of diaphragm Costal margins move out during inspiration.
Line 2. High position of diaphragm. Normal outward movement of costal margins accentuated.
Line 3. Low position of diaphragm. Costal margins move in during inspiration.
Line 4. Very low position of diaphragm. Costal margins move out during inspiration.
Line 5. Actual line of traction of diaphragm. (From T. Wingate Todd.)

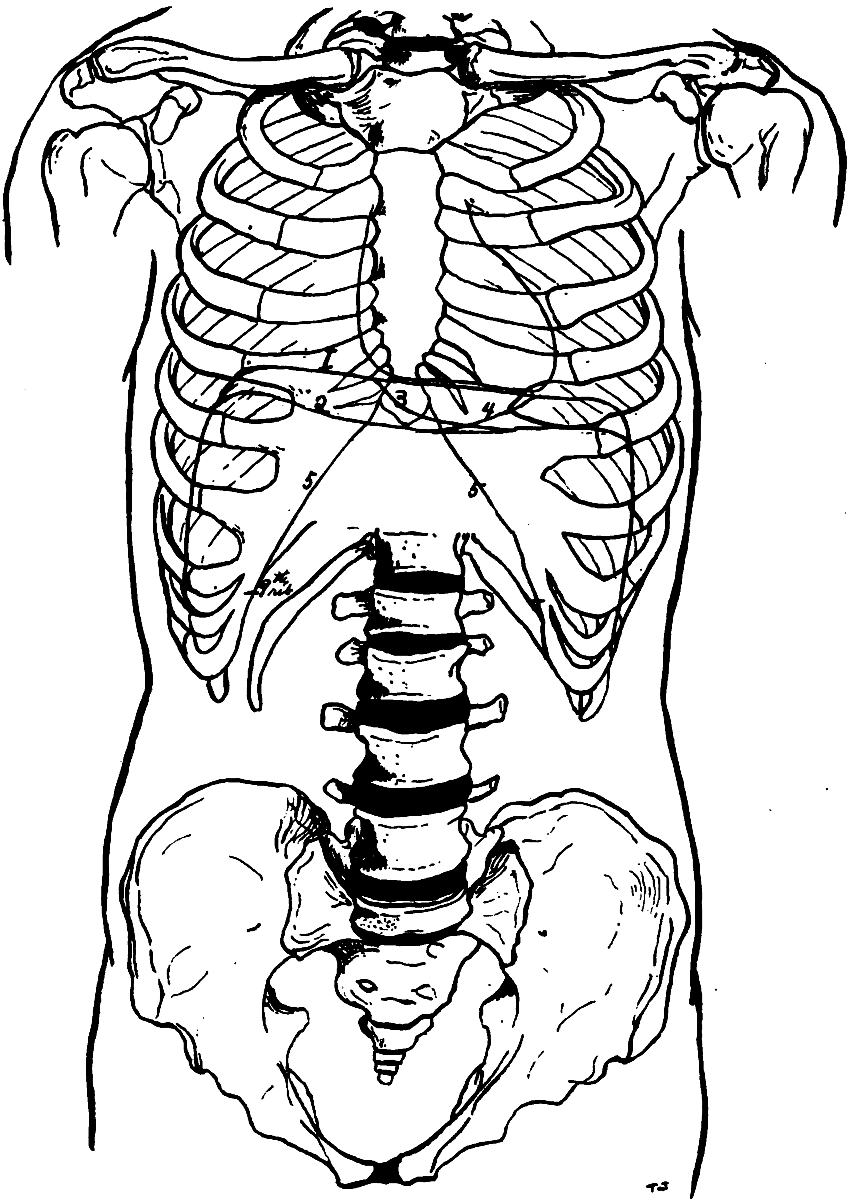


Fig. 117.—Diagram to show the effect of clinical displacements of the diaphragm on the costal angle.

Line 1. Normal position of diaphragm. Costal margains move out during inspiration.
2. Position of diaphragm in general cardiac enlargement. Costal margin from ensiform to ninth rib moves toward median line.
3. Position of diaphragm in left-sided cardiac enlargement. Left costal margin is fixed or moves in during inspiration.
4. Position of diaphragm in right-sided cardiac enlargement. Right costal margin is fixed or moves in during inspiration.
5. Costal margin. (From T. Wingate Todd.)

ternal intercostals have the mastery and cause the costal borders to spread. When the arch of the diaphragm is depressed, as in pleurisy with effusion, emphysema, and empyema, the line of traction and the line of the muscular fibers of the diaphragm correspond more closely, so that the diaphragm is able to use its full force against the intercostal muscles, with the result that the costal border moves towards the median line. The curves of the different fibers of the diaphragm vary greatly; the arch is much less marked in the portion attached to the costal margin near the median line than in that attached in the axillary line. For this reason the anterolateral part of the diaphragm requires less depression to give it a horizontal position than is required for parts occupying a more lateral position. A small pericardial effusion or an increase in the size of the heart may therefore depress the diaphragm sufficiently to give it mastery over the intercostals in the front portion, so that the costal border may here move towards the midline, while the lower borders move in a perfectly normal manner (see Figs. 116 and 117).

During *forced breathing* several muscles are brought into play, among the most important of which are the scaleni, sternomastoid, trapezius, pectorals, rhomboids, and serratus magnus.

There has been considerable debate as to whether *expiration* is normally an active or a passive process. Undoubtedly the expiratory phase under normal conditions does not require the same muscular effort as does that of inspiration, but there are many observations which indicate that expiration is partly under muscular control. The abdominal musculature, for example, increases in tone during expiration, so as to bring about a rise in the abdominal pressure, with the result that the relaxed diaphragm is pushed up into the thoracic cavity. To this extent at least, expiration is accompanied by increased muscular activity.

Before leaving the subject of the diaphragmatic movements, reference must be made to the recent observations of Lee, Guenther and Meleney[3] bearing on the general physiologic properties of the diaphragmatic muscle. They point out that most skeletal muscles in the living body contract with varying degrees of intensity and at irregular intervals, between which relatively long periods of rest occur, but the diaphragm from birth to death performs a continuous succession of brief contractions of fairly regular rhythm and uniform extent, alternating with brief periods of rest. Its muscle fibers, together with those of the other respiratory muscles, therefore hold a unique position among skeletal muscles, which suggests a crude analogy with that of the heart. They have compared the physiologic properties of the diaphragm with those of the extensor longus digitorum, the sartorius, and the soleus, and found

that the diaphragm is composed of a much more efficient muscular tissue than that of the other muscles.

The Effects of the Respiratory Movements on the Lungs.—The changes produced in the dimensions of the lungs by the inspiratory expansion of the thoracic cavity are not uniform, since different parts of these structures are not equally extensible. From an anatomic standpoint, the lungs may be divided into three zones: (1) The inner or root zone containing the bronchus, artery and vein, and their main subdivisions. The large amount of fibrous tissue in this region offers great resistance to any expanding force. (2) The intermediate zone, containing the vascular and bronchial ramifications radiating towards the surface of the lungs, with pulmonary tissue implanted between the rays. This part of the lungs has varying degrees of extensibility, the pulmonary tissue having the most and the vascular and bronchial the least. (3) The outer zone, perhaps 25 to 30 mm. in depth, composed of pulmonary tissue and equally extensible throughout (Keith[1]). The expansion of the lung is accomplished by a moving apart of the less extensible rays of tissue so as to permit the expansion of the more extensible pulmonary tissue between them. Keith compares the mechanism to that seen in the opening of a Japanese fan.

Because the lung expands in the direction of least resistance, study of the inflated dead lung does not reveal the normal expansion brought about by the thoracic movements. In the living body expansion is more limited in some regions than in others. Of the five areas which may be distinguished on the surface of the lungs, three are in contact with relatively immovable parts of the chest wall, and therefore can not be expanded directly. These are: the *mediastinal*, in contact with the pericardium and the structures of the mediastinum; the *dorsal surface*, in contact with the spinal column and the posterior aspect of the thoracic cage, and the *apical surface*. The motions of the first pair of ribs and the manubrium expand chiefly the anterior and ventrolateral part of the apex of the lung, and have only an indirect influence on the dorsal part of the apex—i. e., the part lying directly in front of the necks of the first and second ribs, the most common site of pulmonary tuberculosis. The two surfaces of the lungs which are directly expanded are the diaphragmatic and the ventrolateral or sternocostal. Meltzer[4] found that the negative pressure in the thorax during inspiration was least along the relatively stationary walls of the thorax, and greatest in the regions nearest the diaphragm. From this he concludes that some of the expanding force is lost as it passes through the lungs to the surfaces of indirect expansion. Many observers have claimed that the expansion of the lung does not take place throughout instantaneously and equally. This is illustrated

by the fact that, in the region immediately surrounding a localized consolidation, a fluid has increased resonance, which would not be the case if the relaxation produced was equally distributed throughout the lung.

The root of the lung, which has generally been regarded as more or less fixed, undergoes in normal breathing a definite forward, downward and outward movement, and the heart shares in this movement (Keith). The movements of the lower ribs and diaphragm are responsible for the expansion of the lower lobes and dorsal portion of the upper lobes of the lungs, whereas the movement of the upper five ribs expands the anterior portion of the upper lobes. The relative infrequency of pleuritic friction-sounds and pain over the upper lobes as compared with their frequency over the lower lobes is explained by the fact that the expansion of the upper lobes is accomplished with little displacement of the pleural surfaces, whereas in the lower lobes expansion is accompanied by a gliding of the lungs across the ribs.

CHAPTER XXXVII

THE CONTROL OF THE RESPIRATION

The participation of such widespread groups of muscles in the respiratory act demands that some mechanism be provided to insure its adequate control. With every inspiration, for example, the muscles of the alæ nasi act so as to cause dilatation of the nares, the vocal cords are abducted, and the intercostal muscles, along with the scalenes and the diaphragm are contracting while the muscles of the abdominal wall are relaxing; and all these events occur at exactly the proper time so as to bring about the most efficient opening up of the thoracic cavity. Evidently there must be some mechanism to insure this perfect control. This is effected through the nervous system.

THE RESPIRATORY NERVE CENTERS

The efferent fibers to the various groups of muscle originate in their respective motor neurons, which in most cases are situated in the gray matter of the spinal cord. The harmonious action of these motor neurons, or subsidiary centers, is brought about by the transmission to them of impulses from a higher or master center placed in the medulla oblongata, the pathway of transmission between this master center and the subsidiary centers being in the lateral columns of the spinal cord.

The evidence that *the chief respiratory center* is in the medulla is furnished by observing the effects produced on the respiratory movements by serial destruction of the cerebrospinal axis from above downward. By this method the approximate position of the center is found, its exact location being then determined by punctiform destruction or stimulation of the supposed locus of the center. If we destroy the cerebrum from before backward, piece by piece, we shall find that no marked effect is produced on the respirations until we arrive at about the middle of the medulla, when immediate paralysis of the respiratory movements occurs. If we now proceed to puncture various areas on the floor of the fourth ventricle in another animal, we shall find an area called the *noeud vital,* located about the tip of the calamus scriptorius, destruction of which causes immediate cessation of respiration. It is believed that the center resides in the group of nerve cells known to neurologists as the fasciculus solitarius. It is bilateral.

The subsidiary centers are entirely dependent upon the master center for their harmonious action, as is shown by the fact that if the phrenic motor neuron—which is situated in the cervical spinal cord between the fourth and sixth spinal segments—is isolated from the medulla by a lateral hemisection of the cord just above the fourth segment and by mesial section of the cord opposite the center, the corresponding half of the diaphragm no longer participates in the inspiratory act (see Fig. 118).

The chief center on either side of the midline of the medulla is connected with the motor neurons of *both* sides of the spinal cord, as is proved by the following experiment. When the central end of the vagus nerve is stimulated, the respiratory center becomes excited and the respirations more pronounced, the participation of the muscles on both sides of the body being equal in extent. If now we bisect the medulla down the

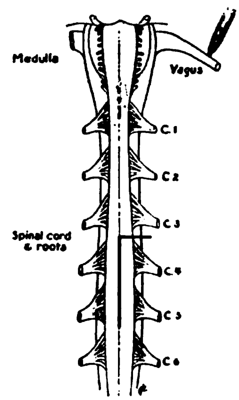


Fig. 118.—Diagram to show cuts required for isolation of the phrenic center.

midline and repeat the stimulation of one vagus, the muscles on both sides will still participate in the increased respiration, which they will likewise do if the cervical cord is bisected or hemisected but the medulla left intact (Fig. 119). The simplest interpretation of these results is that commissural fibers connect both halves of the respiratory center in the medulla and that each half is also connected with the motor neurons of both sides of the spinal cord. Often, especially in young animals, a hemisection of the cord causes cessation of the movements of the diaphragm on the same side; but this paralyzed side at once begins to contract again when the phrenic of the opposite side is cut, probably because the respiratory impulse descending from the chief center, on finding its way along the motor center of the same side of the cord blocked, is forced to follow the crossed path. The crossing in the cord is believed to take place at the same level as that at which the subsidiary center is located (W. T. Porter[12]).

The question now arises as to *how the chief center functionates.* Is it purely reflex in the sense that it depends for its activity entirely on the transmission to it of nervous impulses from elsewhere, or is it automatic in the sense that it can work independently of such impulses? The automaticity of the heart makes it seem not improbable that the center which controls the co-ordinate action of the respiratory muscles would also have an inherent or automatic power. The activity of such an automatic respiratory center would, of course, be subject to great variation as a result of changes in the composition of the blood supplying it, and the fact that it was automatic would not remove it from the influence of nervous impulses. Indeed it is possible to conceive of the automaticity of the center as being of a low order, with its normal functioning dependent upon afferent nerve impulses. Its automaticity might, then,

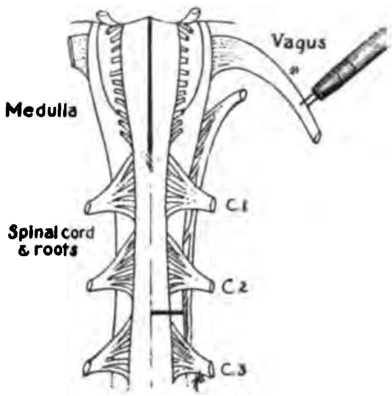


Fig. 119.—Diagram to show certain positions in the medulla and upper cervical cord, where sections may be made without seriously disturbing the respirations. Sections made separately will not disturb the respiration, nor interfere with the effect of vagus stimulation. If both sections are made at once, however, breathing will be seriously interfered with on the side of the hemisection, and this side will not respond to vagus stimulation.

be merely a factor of safety called into play only when the influences ordinarily controlling the center were for some reason removed.

The question which at present confronts us, however, is whether the center may or may not act automatically. Many experiments have been undertaken to test this point, the nature of all of them depending upon the isolation of the center as completely as possible from afferent nerve paths. The most successful experiment has been performed as follows: The influence of the higher nerve centers was removed by cutting across the peduncles of the cerebrum or the pons. The influence of afferent impulses traveling up the spinal cord was removed by completely severing the spinal cord below the level of the phrenic nerves and sectioning all the posterior or sensory spinal roots of the cervical cord above the level of this section. The vagi were also cut to remove the impulses traveling

by them to the respiratory center. By such an operation the only lower respiratory neurons left intact are those of the phrenic nerve, so that the respiratory movements that alone are possible are those in which the diaphragm participates and the muscles of the alæ nasi and larynx. It was found that the animal after the operation went on respiring, though imperfectly, and that the respirations soon became more marked and asphyxial in character, indicating that the blood was not becoming

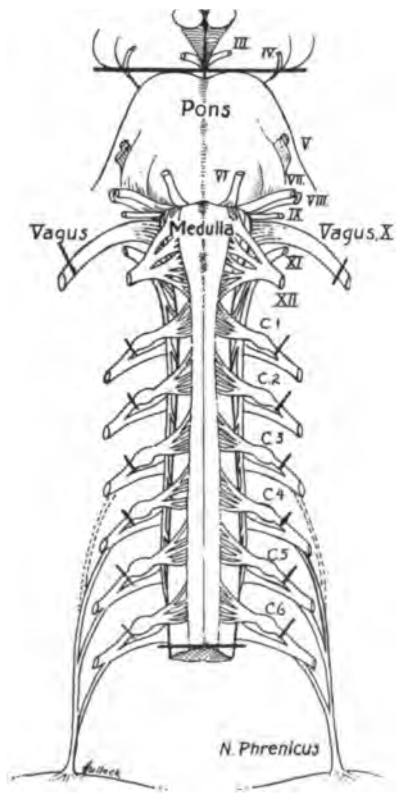


Fig. 120.—Diagram to show where cuts are made to isolate the chief respiratory center from afferent impulses.

properly aerated and that the chemical changes occurring in it were acting directly on the center, stimulating it to greater activity. The conclusion seems warranted that the respiratory center can act automatically, for the only possible afferent nerves left in the above preparation were those carried to the center by the fifth nerve (Fig. 120).

That the respiratory center is extraordinarily sensitive to changes in the composition of the blood flowing through it is a fact that has been known for a long time, but it is only within recent years that the exact

nature of this control and the remarkable sensitivity of the center towards it have been thoroughly established. We shall return to this important subject later. Meanwhile we shall proceed to examine the manner in which the center is affected by sensory impulses transmitted to it.

THE REFLEX CONTROL OF THE RESPIRATORY CENTER

The afferent nerve fibers going to the respiratory center may conveniently be divided into two groups: those which act on it only occasionally, and those which act on it more or less continuously.

The Occasionally Acting Impulses

To the first group belong afferent nerves from practically every part of the body. That impressions from the skin affect the respiratory center is well known by the increased breathing caused by applications of cold water. The influence of these afferent impulses is often very marked, and is frequently taken advantage of in stimulating a newborn infant to take the first breath. Stimulation of the terminations of the fifth nerve in the mucous membrane of the nose, as by inhaling a pungent odor, immediately inhibits respiration. To these occasionally acting afferent impulses may be added the impulses that are conveyed to the respiratory center from the higher nerve centers of the cerebrum. These impulses are largely voluntary in nature, and enable us to hold our breath at will. Some of the cerebral impulses are however also involuntary, their existence being seen by observing the respirations of an animal before and after sectioning the pons or peduncles. The respirations for a time at least become distinctly affected, but they later return with perfect regularity. They may become very irregular, however, if the vagi as well as the pons are cut. Other experimental evidence of the existence of cerebral respiratory fibers is furnished by cerebral localization experiments. During stimulation of the cerebral cortex, for example, a marked effect on the respiratory movements is often noted.

Respiratory rhythm, unlike that of the heart, has often to be modified in order that the respiratory mechanism may be used for other purposes than the ventilation of the lungs. This alteration in rhythm may take the form of a mere inhibition, such as the act of swallowing; or the respiration may be altered, as in phonation and singing. More considerable alteration in the expiratory discharge occurs in coughing and sneezing, and still more in the acts of micturition, defecation, and parturition. We must conclude therefore that the rhythmic stimuli sent out from the respiratory center are so weak that stimuli from other sources may instantly inhibit or change their form at any stage of the cycle.

Stimulation of the endings of the glossopharyngeal nerve inhibits respiration, which explains the holding of the breath that occurs in swallowing.

The Continuously Acting Afferent Impulses

The continuously acting afferent impulses are transmitted to the chief respiratory center by the vagi and their branches, the superior laryngeal nerves. If the vagus nerves are cut or their continuity severed by freezing a portion of them, the respiratory movements become markedly slower. Evidently, the vagus nerves in some way hurry up the respiratory movements. Again, if the central end of either vagus is stimulated with the ordinary interrupted faradic current, a profound effect on the respiratory movements is usually observed. This effect is however not strictly predictable. Usually there is a quickening of respiration, and if the stimulus is a strong one, there may be a standstill of the thorax in the inspiratory position. On the other hand, if the central end of the nerve is stimulated with other types of stimuli, as by slow, weak faradic shocks or by the stimulus produced by the closure of an ascending voltaic current, the effect may be to stimulate expiration rather than inspiration. Such results would seem to indicate that the vagus contains *two kinds of afferent fibers to the respiratory center*, one kind stimulating inspiration, the other, stimulating expiration.

Supposing that such fibers exist, the next question is, how do they become stimulated at their terminations in the lungs? The most natural assumption is that the mechanical distention and collapse of the alveoli which occurs with each respiratory act, serves as the stimulus—an hypothesis to which support is offered by the observation that, when air is blown into the lungs so as to distend the alveoli, the animal immediately makes a forced expiratory movement, whereas when the air is sucked out, the thorax assumes the inspiratory position.

Of the many methods that have been employed to produce distention of the alveoli, the best is undoubtedly that recently employed by Haldane and Boothby.[13] The person or animal is made to respire through a tube in which is inserted a three-way stopcock, which communicates either with the outside air or with a side-tube leading to a spirometer or bag containing air under slight pressure, so that when the stopcock is turned breathing takes place against a definite positive pressure. Such a method is obviously much more physiologic than one in which the air-tube is suddenly clamped at the end of inspiration and the lungs left in a distended condition.

The term used to designate the cessation of breathing is called *apnea*. The extent to which it occurs varies very considerably in different an-

imals and, in the case of man, in different individuals. Thus, when a man is made suddenly to breathe into compressed air, the apnea often lasts for about half a minute, the pause being then broken by a deep expiration followed by a further pause, then again an expiration, and so on with progressively shorter pauses. Disregarding for the present any influences which changes in the composition of the air in the lungs or of the gases in the blood might have in producing the apnea, we may consider the possibility that it is the result of afferent fibers in the vagus. This is an old view, but the most recent experimental evidence does not lend support to it. It was shown by Boothby and Berry,[14] for example, that a similar apnea, though indeed of shorter duration, could be produced in dogs in which the pulmonary branches of both vagus nerves had been severed two months previously. The apnea is, therefore, not a reflex of the vagus, and must be interpreted as due to nervous impulses passing to the respiratory center from some other part of the nervous system, perhaps from centers higher up, or to stimuli transmitted to the respiratory center possibly through afferent fibers in the respiratory muscles.

The formerly very popular theory that respiration is controlled automatically by alternate distention and collapse of the alveoli, acting through the afferent fibers of the vagus nerve on the respiratory center in such a way as to bring the opposite act with each expiration and inspiration, must, therefore, be abandoned. But we can not deny that the vagus plays a most important role in the control of the function of the respiratory center, for apart from the effect which we have seen to follow the severence of continuity of the nerve, there is the important observation of Alcock and others[15] that when nonpolarizable electrodes are placed on the vagus nerve and connected with a galvanometer, a current of action occurs toward the end of each inspiration in quiet breathing; and when the respirations are forced, a current of action appears during both inspiration and expiration. Another reason for believing that the vagi have some important function to perform in connection with the control of respiration is the fact, observed by F. H. Scott,[16] that in an intact animal, when atmospheres containing increasing percentages of carbon dioxide are respired, the respirations become both deeper and quicker, whereas in one whose vagi have been cut the carbon dioxide causes only a deepening of the respirations. From this result it would appear that the vagi exert an influence on the rate of the respirations but not on their depth, this effect, as we shall see later, being dependent primarily on changes in the composition of the blood supplying the respiratory center. It is probable that both controlling agencies act together, the one serving to maintain the center in a proper state of

excitability, and being active to a greater or less extent all the time; while the other acts only occasionally on the "tuned up" center. There is, of course, no doubt that it is through the nerves that the occasional alterations of respiration occur. They appear also to have a certain influence on the rhythm, for Stewart, Pike and Guthrie[17] observed that, after resuscitation from acute brain anemia, the respirations when they returned were of the same rhythm as that of the artificial respirations employed during the resuscitation.

The usually accepted hypothesis as to the mechanism by which the nerve impulses hasten the respiratory movements is that an afferent impulse is transmitted to the respiratory center towards the end of each inspiration, which has the effect of inhibiting the inspiratory discharge from the center and thus cutting short the act of inspiration so that expiration automatically supervenes. This explanation is in agreement with the fact that quiet inspiration involves activity on the part of the respiratory muscles, whereas expiration is usually almost entirely passive, being due to the return to a resting position of the stretched and displaced structures. On the other hand, in forced respiration and in certain animals under normal conditions, expiration becomes active, in which event a current of action becomes evident in the vagus nerve during the expiratory phase.

The superior laryngeal branch of the vagus should really be classified as one of those nerves that have an occasional influence on the respiratory center, its particular function being in connection with the act of coughing. When a foreign body irritates the mucous membrane of the larynx, the nerve fibers transmit impulses to the respiratory center which excite a violent expiration and at the same time cause the glottis to close. The closure of the glottis lasts, however, only during the first part of the expiration; it then opens, with the result that the sudden release of intrapulmonic pressure causes the expulsion of the foreign substance in the air passages.

CHAPTER XXXVIII

THE CONTROL OF RESPIRATION (Cont'd)

THE HORMONE CONTROL OF THE RESPIRATORY CENTER

Just as the rhythmical activity of the heart is readily influenced by changes in the composition of the blood supplying it, so also is that of the respiratory center. In the case of the heart it is the cations—calcium, potassium and sodium—that have the most pronounced effect, whereas in the case of the respiratory center it is largely the relative concentration of hydrogen and hydroxyl ions—the H-ion concentration (C_H) of the blood. This influence can be shown in a general way by injecting acid or alkaline solutions into the peripheral end of the carotid artery of an anesthetized animal, or better still of one that has been decerebrated. Acid injections stimulate the respiratory activity; alkaline injections tend to depress it. When the acid or alkaline solutions are injected intravenously in other parts of the body, so that they become thoroughly mixed with the blood before the respiratory center is reached, the effects are not nearly so pronounced, because the buffer influence of the blood has time to develop (see page 36).

From the results of such injection experiments, however, one could not draw the conclusion that under *normal* conditions the activity of the respiratory center is affected by measurable changes in C_H of the blood, for, as we have seen, constancy of C_H is one of the most remarkable properties of the animal fluids. To justify the conclusion that the respiratory center is affected by changes in C_H, it is necessary to observe the behavior of some easily measurable acid or alkaline constituent of the blood that undergoes changes in amount that are proportional to an alteration in C_H. In order to understand what this acid or basic substance may be, it will be advisable to recapitulate the main factors concerned in maintaining C_H at a constant level. This value is obviously dependent upon the balance between basic and acid substances, so that any variations which it undergoes must be caused by changes in the relative amount of one of these. Changes in *base* may occur, exogenously, by altering the alkali content of the food, or, endogenously, in various ways but particularly by variations in the amount of ammonia produced during the course of metabolism of protein. Thus, when sudden demands are made by the organism for an increased amount of base,

the amino groups—split off from the amino bodies—become converted into ammonia instead of into the neutral substance, urea. But the chief variations seem to concern *acids* rather than the basic substances. These acids may be divided into three groups: *fixed inorganic acids,* represented by phosphoric; *fixed organic acids,* represented by lactic; and *volatile acids,* represented by carbon dioxide. Of these three groups, the first shows the least tendency to change, and the third, the greatest. Changes in the second group (fixed organic acids) are effected partly by excretion through the urine and partly by oxidation into volatile acid. The sudden and rapid changes in the third group are brought about by the diffusion of the CO_2 of the blood into the alveolar air. Gross changes in the acid content of the blood are therefore mainly effected through alteration in the excretion of the fixed acids, whereas sudden changes are effected by excretion of the volatile acid. It is important to note here that the fixed organic acids do not participate to any great extent in the makeup of the acid content of normal blood: they appear only under unusual conditions, as in dyspnea. The variations in C_H that ordinarily affect the activity of the respiratory center are therefore dependent upon changes in the volatile acid, a direct measure of which is found in the tension of CO_2 in the blood. The correlation between C_H of the blood and respiratory activity must be a very close one if C_H is to be maintained.

The Laws of Gases.—In order to understand the principles upon which alterations in CO_2 tension are dependent, it will be necessary for us to review briefly some of *the gas laws.* Among these laws the first in importance is the *law of pressure,* which states that, other things being equal, the pressure of a gas is inversely proportional to its volume; if a gas occupying a certain volume is compressed by a pump so that it occupies one-half of its previous volume, its pressure will become doubled. The second is the *law of partial pressure,* which states that the partial pressure of a gas in a mixture of gases, having no action on one another, is equal to that which this particular gas would exert did it alone occupy the space occupied by the mixture. Thus, atmospheric air consists roughly of 79 volumes per cent of nitrogen and 21 of oxygen; the partial pressure of the oxygen is therefore equal to $\frac{21}{100} \times 760$ mm. Hg, this last figure being the barometric pressure of air at sea level. The third is the law of *solution of gases,* which is to the effect that the amount of gas which goes into solution in a liquid having no chemical attraction for the gas, is proportional to the partial pressure of gas. If water is exposed to air, the amount of oxygen which it dissolves will be the same as if the water had been exposed to oxygen at a pressure equal to that

of the partial pressure which it produces in air. The same will be the case with the nitrogen of the air. The actual amount of gas which becomes dissolved in the fluid, pressure and temperature being constant, depends partly on the nature of the gas and partly on the nature of the fluid. For example, the solubility of oxygen in water is considerably different from that in a neutral oil; or, taking the same solvent, nitrogen and CO_2 do not dissolve to the same extent in water. It becomes necessary, therefore, in calculating what amount of a particular gas will dissolve in a particular fluid to use a figure known as the *coefficient of solubility* of the gas—that is, the amount of gas taken up by a unit volume of fluid at standard temperature and pressure; for example, to say that the coefficient of absorption of nitrogen in water at 0° C. is 0.0239 means that, at this temperature and at normal barometric pressure, 1 c.c. of water will dissolve 0.0239 c.c. of nitrogen when exposed to a pure atmosphere of this gas. Obviously, then, if water were exposed to 79 per cent of an atmosphere of nitrogen (as in air) the amount which would become dissolved in each c.c. would be $\frac{79}{100} \times 0.0239 = 0.0189$ c.c.

In solutions containing no chemical substances with which the gas can enter into combination, it is evident that the tension of the gas will be proportional to the amount of gas that can be displaced or pumped out from the fluid. On the other hand, when a chemical compound is formed, the combined gas will exercise no direct influence on the tension, so that this will be independent of the amount; *in such cases separate methods will have to be used for the determination of amount and tension.* Let us take the case of pure water exposed to an atmosphere of CO_2: the amount of CO_2 which goes into solution will depend entirely on the pressure. If a trace of alkali is dissolved in the water, however, some of the CO_2 will become combined to form carbonate, so that a much larger quantity of CO_2 will be displaceable from the solution (as by adding a mineral acid to it) than corresponds to the tension of CO_2 in the atmosphere surrounding it. Since blood contains alkali the conditions are analogous with those of a weak alkaline solution.

The Tension of CO_2 and O_2 in the Arterial Blood.—If we were to pass blood at body temperature in a very thin film over the walls of a confined space containing a mixture of gases one of which was CO_2, it is evident that the percentage of CO_2 in the atmosphere contained in this space would remain unchanged only when the tension of this gas in the blood was the same as that in the confined atmosphere. If, on the other hand, the tension of CO_2 in the blood should correspond to a percentage that is higher than that in the atmosphere, then CO_2 would diffuse from the blood, and at the end of the experiment an analysis of the

atmosphere in the space would show that the CO_2 percentage had been raised. If the blood contained a lower tension than that corresponding to the percentage of CO_2 in the space, some of the CO_2 would diffuse into the blood, and its percentage in the atmosphere would be lowered. By successively exposing blood to gas mixtures that contain slightly different percentages of CO_2, we should ultimately find one with which the free CO_2 in the blood was in perfect equilibrium, and we should be able to state that the tension of this gas in the blood was equal to a certain percentage in the atmosphere surrounding the blood (see Fig. 121).

Many forms of apparatus based on the above principle have been invented for the examination of the tension of the gases in the blood. The most accurate is that devised by Krogh,[18] the principle of which

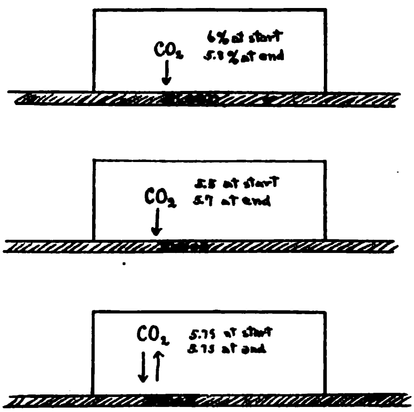


Fig. 121.—Diagram to show principle for measurement of the tension of CO_2 in blood. The CO_2 tension of blood is supposed to be 5.75.

differs slightly from that just described in that a bubble of air is exposed to a relatively large quantity of blood, so that after a time actual equilibrium of gas tension becomes established between the bubble and the gases of the blood. This apparatus is shown in Figs. 122 and 123. It consists of a graduated tube of narrow bore surrounded by a water jacket. To the upper end of the graduated tube a small syringe is attached. The lower end of the graduated tube expands into a thistle-shaped bulb, closed below by a cork, through which is inserted a tube (inflow tube) ending near the top of the bulb in a fine opening and connected outside with an artery. An outflow tube is also connected with the thistle-shaped bulb.

At the beginning of the experiment the thistle-shaped bulb and the graduated tube are filled with physiologic saline. By means of the syringe a small bubble of air is then introduced, so that it lies at the

junction of the thistle-shaped bulb and the graduated tube. As the blood is allowed to enter through the inflow tube, it is ejected in a fine stream around the bubble of air, which moves about in the stream. The blood displaces the saline out of the bulb into the side tube. After the bubble has been subjected to the influence of the blood for some minutes, the gases in it come into perfect equilibrium with those in the blood. The percentage of O_2 and CO_2 in the bubble will therefore correspond to the tension of these gases in the blood. The analysis is effected by drawing the bubble into the graduated tube by means of the syringe,

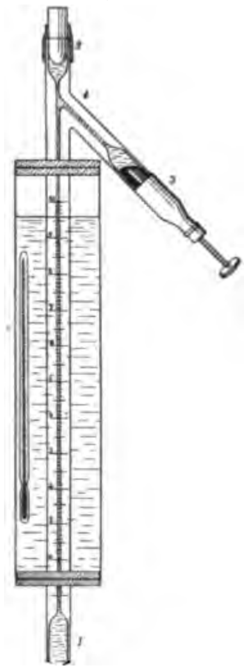


Fig. 122.

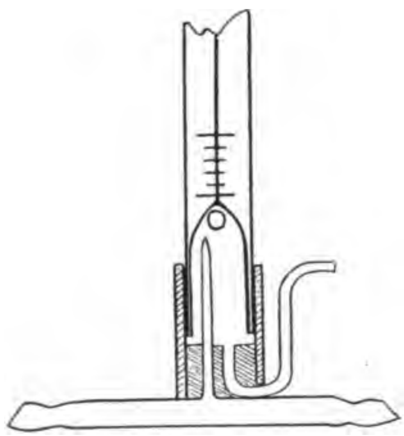


Fig. 123.

Fig. 122.—The gas analysis pipette for the microtonometer shown in Fig. 123. For description see context. (From A. Krogh.)

Fig. 123.—Microtonometer, to be inserted into a blood vessel. The small circle represents the bubble of air. For further description see context. (From A. Krogh.)

measuring its capacity, transferring it into a bulb containing KOH, which absorbs the CO_2, then taking it back into the capillary tube and again measuring. The shrinkage obviously corresponds to the amount of CO_2. The bubble is then transferred into potassium pyrogallate solution, where the O_2 is absorbed.*

The Tension of CO_2 and O_2 in Alveolar Air.—Having seen how we may determine the tension of the gases in blood, we must now consider

*Since the above was written, a more efficient tonometer devised by the late T. G. Brodie has been described by O'Sullivan (Am. Jour. Physiol., Sept., 1918).

the method by which the tensions of these gases in alveolar air can be determined. The simplest and until recently the most accurate method is that of Haldane and Priestley.[19] This consists in having an individual, with his nostrils clamped, breathe quietly through a piece of hose pipe about a meter long, which has at the mouth end a short side-tube leading to an evacuated gas-sampling bulb of about 50 c.c. capacity.* After the subject has become accustomed to breathing through the tube, he is asked to make a forced expiration and at the end of it to close the mouthpiece with his tongue. At this moment the operator opens the tap of the sampling tube, allowing the air from the tubing through which the individual has made the forced expiration to rush in and fill it. This sample represents the air from the alveoli (see page 302), and is analyzed for percentages of CO_2 and O_2. Since each normal inspiration dilutes the alveolar air somewhat, it is necessary, for constant re-

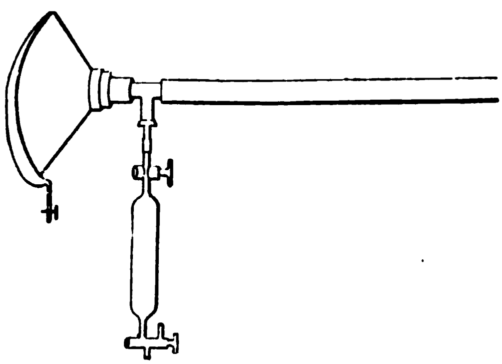


Fig. 124.—Apparatus for collection of a sample of alveolar air by Haldane's method. It is better to use a mouthpiece than a mask.

sults, to make two analyses of alveolar air from each subject, one taken at the end of a normal inspiration and the other at the end of normal expiration. The average of the two results is taken as the composition of the alveolar air.

On account of the difficulty in securing intelligent cooperation in the application of this method, particularly with children, others have been devised. One of the simplest is that of Fridericia, which is a modification of the Haldane-Priestley method, the apparatus for which is shown in the figure (Fig. 125), and the manipulation of which is outlined in the legend. Another is to take a mixed sample of the very last portion of several normal expirations. On account of the extended use which is being made of measurements of alveolar air composition, both in lab-

*In place of the gas-sampling tube it is much more convenient and equally accurate to employ one of the modern ground glass piston syringes (Luer). The piston should, of course, be well smeared with a good mineral grease.

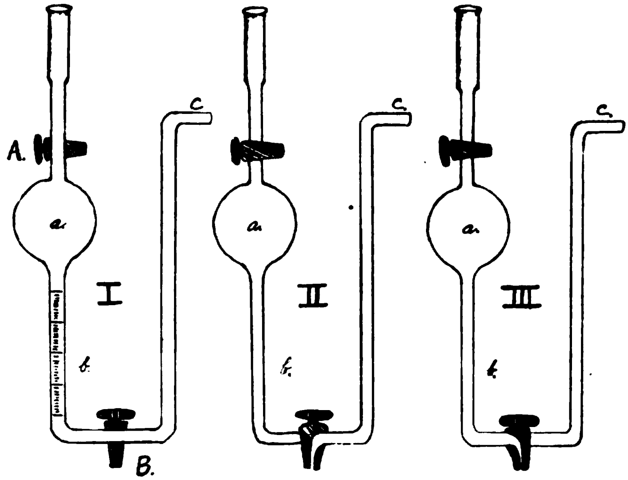


Fig. 125.—Fridericia's apparatus for measuring the CO_2 in alveolar air. The person expires forcibly through the tube with the stopcocks as in I. *A* is closed and the tube placed in water to cool the air, after which *B* is turned as in II. The entrapped column of air equals 100 c.c. A solution of caustic alkali is now sucked into *C* with stopcocks as in II. *B* is then turned as in I but with *A* still closed, and the alkali solution allowed to enter *b*, after which *B* is turned off, the excess of alkali solution in *C* allowed to run out and the burette shaken. The burette is then submersed up to *a* in a cylinder of water, with *B* as in III. After allowing for cooling, the level at which the water stands gives the per cent of CO_2.

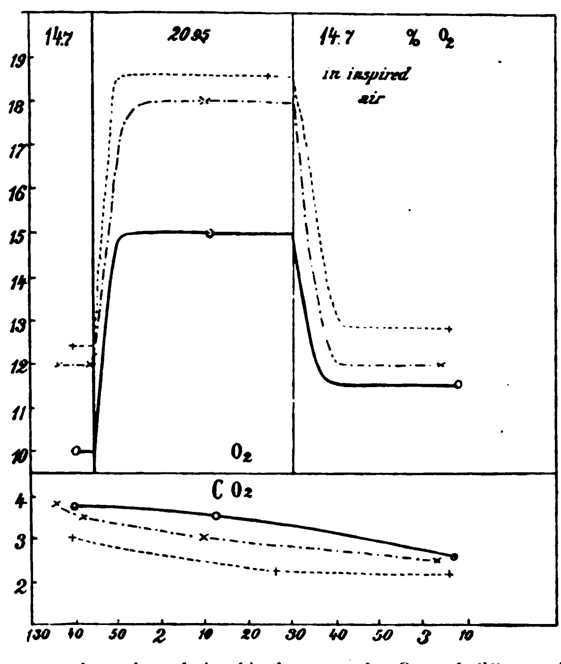


Fig. 126.—Curves to show the relationship between the O_2 and CO_2 tensions in alveolar air (dotted lines) and arterial blood (continuous lines). It will be observed that the tension of CO_2 in blood is slightly above that in alveolar air, but that the reverse relationship obtains for O_2. In the upper part of the curve the O_2 tension in the alveolar air was experimentally altered, causing a corresponding alteration in the O_2 tension of the blood. This result is of practical significance in connection with O_2 alterations in gas poisoning, pneumonia, etc. (From A. and M. Krogh.)

oratory and in clinical work, a special chapter has been devoted to the subject, giving in detail the more recent methods devised by R. G. Pearce.

Lastly, it should be noted that several observers believe that a more reliable estimate of the alveolar tension of CO_2 (and of O_2) can be made by analyzing a sample of ordinary expired air and calculating the percentages of CO_2 and O_2 in the alveolar air by allowing a constant dead-space capacity of 140 c.c. (Krogh, etc.).

If we compare the CO_2 tension of arterial blood, as measured by the Krogh method, with that of alveolar air, we shall find that there is a remarkable correspondence, indicating, therefore, that, when the arterial

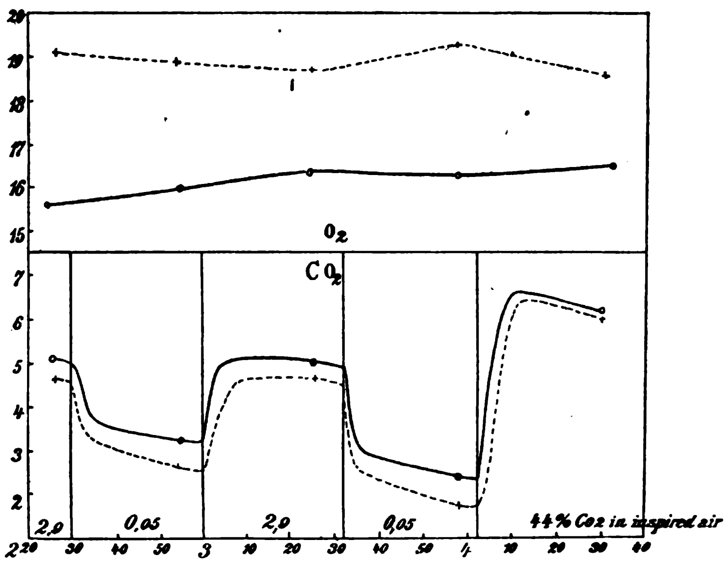


Fig. 127.—Same as Fig. 126, except that in this case the tension of CO_2 in the alveolar air was experimentally altered. (From A. and M. Krogh.)

blood leaves the alveoli, its partial pressure or tension of CO_2 is exactly equal to that in the alveolar air. This is shown in the accompanying curves of experiments performed by Krogh. The dotted line in these curves represents the tension of CO_2 or O_2 in alveolar air, and the continuous line, these tensions in arterial blood. Close correspondence will be observed between the CO_2 curves even when sudden changes in alveolar CO_2 were induced by artificial means. In the case of the O_2 tensions, however, that of the blood is always lower than that of the alveolar air, the differences being especially marked when the O_2 tension in the alveoli is raised (Figs. 126 and 127).

Tension of CO_2 in Venous Blood.—If we examine the CO_2 tension of the venous blood coming to the lungs, we shall find that it is distinctly

higher than that in the alveolar air. The earliest method for measuring it consisted in passing a lung catheter into the right bronchus and then blocking the passage above the open end of the catheter by inflating a rubber collar or ampulla. The renewal of air in the right lung is thereby prevented, and a sample of the stagnant air can be removed and analyzed. In such a case, however, the blood will have circulated several times round the body, and with only one lung operating the risk is incurred that more CO_2 is being discharged into the blocked lung than corresponds to the tension of CO_2 of venous blood under normal conditions.

Much more practical methods are those of Haldane, Yandell Henderson and R. G. Pearce, which are much the same in principle. In Pearce's method, the person first of all inspires from a gas meter containing a gaseous mixture with about 10 per cent of CO_2. Immediately after filling the lungs, he makes a rapid forced expiration into a tube provided with a valve having four openings. This valve is turned through a complete circuit during the expiration, so that four fractions of the expired air can be collected in rubber bags connected with side tubes opening opposite the four openings in the valve. The first fraction will contain a little less than 10 per cent CO_2, the second distinctly less, while the fourth will contain the same as the third, indicating that equilibrium between the CO_2 of the alveolar air and the blood must have been attained. This figure therefore gives us the tension of CO_2 in the venous blood of the lungs. In Henderson's method the rebreathing is performed into gas receivers containing 6 per cent CO_2.

These results then indicate that the whole process by which CO_2 is exchanged in the lungs is dependent on the law of gas diffusion; the gas diffuses from a place of lower to a place of higher pressure, and does so until equilibrium is attained.

CHAPTER XXXIX

THE CONTROL OF RESPIRATION (Cont'd)

THE ESTIMATION OF ALVEOLAR GASES

BY R. G. PEARCE, B.A., M.D.

Methods such as that of Haldane and Priestley, which calculate the mean percentage composition of the alveolar air by analysis of a sample taken from the end of a prolonged forced expiration, give values which are too high for CO_2 and too low for O_2. There are several reasons for this: (1) In the time taken for the prolonged deep expiration an appreciable amount of CO_2 will be given off by the blood to the alveolar air, and oxygen will be absorbed—that is, the sample will not contain the same percentages of CO_2 and O_2 at different stages of expiration. (2) The portion of the tidal air which reaches the alveoli dilutes the alveolar air and thus causes the amount of CO_2 given off by the blood to vary during the different phases of respiration. If we bear in mind that the tensions of CO_2 in the alveolar air and in the blood leaving the lungs are always the same (page 343), and that the entire fall in CO_2 tension in the alveolar air occurs during inspiration, then it is clear that the blood in the pulmonary capillaries must have a maximum tension and load of CO_2 at the end of expiration, and a minimum tension and load of CO_2 at the end of inspiration. Accordingly, the average of the percentage of CO_2 and O_2 at the end of inspiration and expiration, as determined by the Haldane-Priestley method or by any of its modifications, must fail to give the correct mean tension of these gases in the alveolar air during expiration. The error which makes the CO_2 higher than it should be, makes the percentage of O_2 less than it should be. These influences taken along with the fact, which will be shown later, that the evolution of CO_2 from the blood is relatively more rapid at low than at high tension of CO_2, indicates that the blood in the pulmonary capillaries during inspiration must contribute a greater part of the CO_2 excreted during a respiratory cycle than that in the pulmonary capillaries during expiration, and moreover that a greater part of the CO_2 excreted must be evolved at a tension which is below the mean tension of the CO_2 present in the entire time of the expiration. We conclude, therefore, that the average tension of CO_2 in the alveolar air, determined

by the actual tension under which the gas is evolved from the blood, is less than the average tension of CO_2 in the alveolar air during the time of a respiratory cycle.

In the case of O_2 the conditions are different. While the diluting effect of the alveolar tidal air is marked in altering the amount of CO_2 given off during the different phases of a respiration, it can have little influence on the amount of O_2 taken up by the blood under normal conditions. This is evident from a study of the dissociation curve of hemoglobin (page 383), which shows that at tensions above 65 mm. Hg the hemoglobin is practically saturated with O_2. Since the tension of O_2 in the alveolar air under normal conditions is greater than 65 mm. (95-100 mm.), the rate of absorption of O_2 must be practically maximal during the respiratory cycle—that is, it will not change at different phases of it.

While the relationship of the alveolar gases is continually changing at different stages of the respiratory cycle, their mean relationship for periods including several respirations or for complete respirations is more or less constant, being controlled by the type of the metabolism, and mathematically expressed by the respiratory quotient (page 547). The average relative percentages of the two gases in the alveolar air must therefore be the same as in the tidal air. In the alveolar air collected by the Haldane method, however, the above factors cause the respiratory quotient to be less than that in the tidal air.

These points have been insisted upon because much of the knowledge of the gaseous exchange between the blood and the air in the lungs, as well as the control of respiration, has been built upon data obtained by the Haldane-Priestly method, and in considering this work, which we shall do in subsequent pages, it is advisable that we be aware of the limitations of the method employed. The method has been an invaluable one for opening up a hitherto entirely unexplored field of research, but now, the pioneer work having been done, we must employ methods which will enable us to explore more exactly.

An Accurate Standard Method for Normal Subjects.—The most accurate method, and one free from many of the theoretic errors present in the others, depends on the relationship found to exist between the diluting effect of the air in the dead space (see page 302) and the known percentage composition of the alveolar air in expirations which are of varying depths but *of equal and normal duration* and which follow normal inspirations (R. G. Pearce).

In this method the subject is made to breathe through valves, which automatically separate the inspired from the expired air. The expired air is led into a tube connected with two spirometers by two three-way stopcocks. The spirometers are of the

Gad-Krogh type, one being capable of holding ten liters, and the other one and a half. The exact time during which air enters is recorded by the small spirometer by means of a grooved dial on the axis of the lid, on which a thread works over a system of pulleys, and any movement is accurately recorded by a writing point on the smoked paper of a drum. The spirometers are connected so that the air current may be directed in the three following ways: (1) through Cocks *1* and *2* outside; (2) directly through both cocks into the large spirometer for the purpose of collecting a series of expirations; and (3) through Cock *1* directly into the small spirometer for catching a single expiration. In all experiments the first filling of the spirometer is rejected, so that the dead space of the spirometers is filled with air of approximately the same composition as in the succeeding expirations. The time is marked in seconds by a time clock. The respiratory movements are recorded by a pneumograph. (Fig. 128.)

The subject is brought into respiratory equilibrium by having him breathe through the valves for a period of time before the observation. The respiratory movements during this time are recorded while the cocks are in Position *1*. When the observation is started, the cocks are turned into Position *2* during the time an inspiration is being

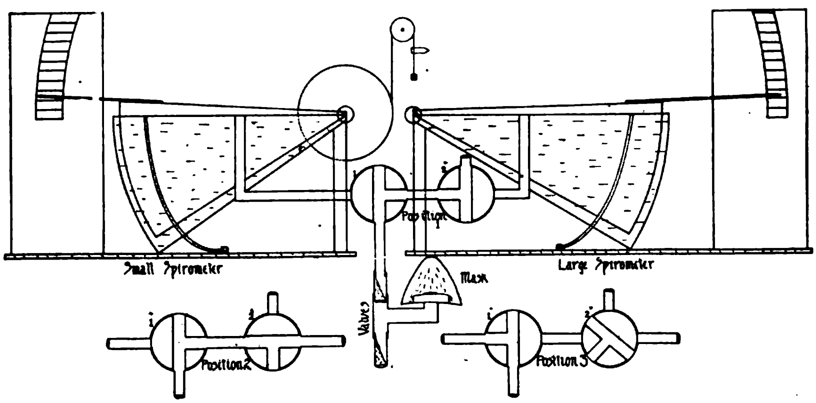


Fig. 128.—Arrangement of meters and connections of Pearce's method for measurement of CO_2 of alveolar air in normal subjects.

made, so that the expirations which follow may be collected in the large spirometer. After about ten respirations (a counted number) have been collected, the cocks are turned to Position *3* during an inspiration, and a single deep expiration is collected in the small spirometer. In order that the time of this may be the same as the normal expiration, it is necessary to quicken it. This is more or less a chance procedure, but with a little training, the operator can close the stopcock with sufficient accuracy to interrupt the deep expiration at the end of the normal expiratory time. Should there be any gross variation from the normal expiratory time, the sample must be collected again. Not infrequently the inspiration immediately preceding the expiration into the small spirometer is varied involuntarily by the subject on account of his being aware that the following expiration has to be deepened and quickened; this can be partially overcome by giving him the signal to breathe out deeply after he has actually begun to expire.

Determinations are made of the average volume of the tidal air (c.c. air in large spirometer divided by number of breaths), of the volume collected from the deep expiration, and of the percentage composition of the tidal air and that of the deep expiration. A criterion for determining whether or not the procedure has been carried

out correctly is the respiratory quotient (ratio of CO_2 excreted to O_2 absorbed). For reasons which are set forth above, the quotients should be approximately equal in the air collected in the large and in the small spirometers; if they are not so, the conditions of the method have not been correctly carried out.

Since the dead space and the *average* composition of the alveolar air under these conditions may be considered constant, the percentage composition of the deep expiration will differ from that of the mixed sample of several normal expirations in proportion as the dead space exerts a greater diluting effect in the small than in the large expiration. This being the case, the data obtained can be combined algebraically to give either the capacity of the air passages or the percentage composition of the alveolar air.

Let A = amount of air in large expiration (small spirometer),

Ai = amount of air in small or normal expiration (tidal air),

B = the percentage of CO_2 or O_2 in the expired air of large expiration,

Bi = the percentage of CO_2 or O_2 in the expired air of small expiration,

x = the capacity of the dead space,

y = the average percentage of CO_2 or O_2 in the alveolar air; then,

$A \times B = (A - x)y$ and $Ai \times Bi = (Ai - x)y$.

Solving this for x, y remaining constant under the same physiologic conditions, we have: $x = \frac{A \times Ai \times (B - Bi)}{A \times B - Ai \times Bi}$, the dead space. Or solving for y, we have:

$y = \frac{A \times B - Ai \times Bi}{A - Ai}$, the mean percentage of CO_2 in the alveolar air. In case the dead space for O_2 is desired, B and Bi must be made to equal the O_2 absorbed.

Clinical Method.—The use of the kymograph and pneumograph, and other complicating factors, make the method as just described quite impracticable for clinical procedure, but the use of the same apparatus with the following modification will yield satisfactory results for most clinical purposes. The patient is made to respire through the valves for a short time, after which the observer collects a single expiration in a small spirometer by turning the stopcock from Position 1 to 2. A sample of this is taken for analysis, and the spirometer is again emptied and a series of successive samples of deeper expirations taken. This is done by directing the patient, after he has started to breathe normally into the spirometer, to breathe more deeply. The amount of air collected in each expiration is controlled by the observer by closing the stopcock when the desired volume is obtained. By this means one can collect several expirations differing from one another by increasing amounts but all occupying the same time. The samples of the various expirations are collected in a series of numbered sampling syringes, and the gaseous composition of each is determined. When the percentage of CO_2 or O_2 in each expiration is plotted on cross section paper on the ordinates, with the volume of the expirations in c.c. on the abscissæ, a hyperbolic curve should be obtained. Any marked deviation from such a curve indicates that some error has been made in taking a sample, and

this observation should be discarded. The different observations are then combined in the formula given on page 347. The determination of the CO_2 percentage of expired air is so simple that a number of specimens of varying depths of expiration can be taken and thus many points on the curve determined. For the most accurate results it is in general best to compare only those expirations which differ from one another by at least 0.3 per cent in CO_2 and by at least 200 c.c. in volume. This depends on the fact that the diluting effect of the dead space in reducing the percentage of CO_2 in the expired air from that in the alveolar air is greater in relatively small expirations. If more exact work is desired, the O_2 content can be determined on each specimen, the respiratory quotient calculated, and only those expirations which show the same respiratory quotient combined.

In the table each observation is compared with each of the others in all possible combinations.

NO. OF OBSERVATION	TOTAL EXPIRED AIR	PER CENT CO_2 IN EXPIRED AIR	ALVEOLAR CO_2			DEAD SPACE		
			1	2	3	1	2	3
1	450	3.10						
2	637	3.66	4.99			170		
3	750	4.00	5.34			189		
4	960	4.28	5.30	5.48	5.27	189	183	214
5	1120	4.30	5.11	5.15	4.92	161	140	184
6	1440	4.40	5.16	4.98	4.82	171	127	171

General average for CO_2 in alveolar air, 5.13.
General average for dead space, 172. Dead space in valves in this experiment was about 30 c.c.

Another method which has been suggested for clinical purposes is that of Plesch; this consists in having the subject breathe several times in and out of a small bag. It is assumed that after such respiration the composition of the air in the bag will become the same as that in the alveoli. Although this is no doubt true, it has been shown that the method is fallacious, because the CO_2 tension determined in this way is not that of the arterial blood alone, but is the average between it and that of the venous blood.

CHAPTER XL

THE CONTROL OF RESPIRATION (Cont'd)

THE NATURE OF THE RESPIRATORY HORMONE

The practical importance of the observations described in the foregoing chapters in the investigation of the relationship between C_H of the blood and respiratory activity will now be plain, and it remains for us to consider the physiologic evidence that such a relationship exists. In the first place, let us consider **the behavior of the acid-base equilibrium during conditions of abnormal breathing—hyperpnea and dyspnea.***

As CO_2 accumulates and O_2 becomes used up in a confined space, the breathing becomes intensified. In searching for the exact cause of this effect, we must first of all ascertain whether the hyperpnea is due to the deficiency of O_2 or to the accumulation of CO_2. Many of the experiments bearing on these problems can be more satisfactorily performed on man than on laboratory animals, because anesthesia is not necessary and the subjective symptoms experienced are of great value in the interpretation of the results. If an individual is placed in a large air-tight chamber (2000 liters' capacity), and the depth and rate of breathing observed as the CO_2 accumulates and the O_2 becomes used up in the air of the chamber, no distinct change in respiration will be observed until the CO_2 percentage of the air has risen to almost 3. Above this point, however, the hyperpnea becomes more and more pronounced, until finally, when the CO_2 percentage has risen to about 6 and the O_2 percentage has fallen to 13.5, it becomes unbearable (dyspnea). From the results of the foregoing observation alone we could not, however, decide whether the excitation of the respiratory center is due to the deficiency of O_2 or to the increase of CO_2. If the experiment is repeated with the difference that the CO_2 as it accumulates is absorbed by soda lime, no hyperpnea will develop even when the O_2 is as low as in the previous experiment. We may conclude, therefore, that in the first experiment *CO_2 accumulation must have acted as the respiratory stimulus.*

The same conclusion is arrived at as a result of observations on individuals caused to breathe in a more confined space as into a rubber bag of about 225 liters' capacity. Under these conditions hyperpnea de-

**Hyperpnea* means slightly increased breathing; *dyspnea*, labored breathing, but yet with sufficient ventilation to maintain life; *asphyxia*, the results of insufficient breathing.

velops more rapidly than in the large cabinet, and a higher percentage (10 per cent) of CO_2 can be tolerated. That in this case also deficiency of O_2 is not responsible for the hyperpnea can be shown by repetition of the experiment either with an excess of O_2 in the bag or with absorption of the CO_2 by soda lime. In the former case hyperpnea will develop as usual, while in the latter it will not supervene until the percentage of O_2 has fallen below 10, when cyanosis becomes marked. In fact, some people become cyanosed and unconscious, and collapse under these conditions before there is any respiratory disturbance. A peculiarity of the effect of O_2 deficiency is that the person may be unaware of the seriousness of his condition; indeed he may be somewhat stimulated. The conclusion may be drawn that *deficiency of O_2 per se can serve as a respiratory stimulus only when it is so extreme as to cause other serious symptoms.* This conclusion does not rule out an important influence of O_2 deficiency in increasing the excitability of the center towards CO_2. Under ordinary conditions, however, the center is far more sensitive towards slight changes in the CO_2 percentage.

There is an obvious reason why the adjustment of pulmonic ventilation should not depend upon changes in O_2 supply to the respiratory center. If it were so, many other tissue activities and other nerve centers would suffer from the O_2 deficiency before there was time for the breathing to become stimulated sufficiently to make good the loss of O_2. As a matter of fact, headache, dizziness, nausea and even fainting are almost certain to be caused whenever any muscular exercise is attempted in an atmosphere containing a deficiency of O_2 but no excess of CO_2 (cf. mountain sickness). An adequate O_2 supply of the body is, therefore, insured by changes in CO_2 tension of the blood.

Quantitative Relationship between CO_2 of Inspired Air and Pulmonary Ventilation.—These results suggest, as the next step in the investigation of our problem, the determination of the quantitative relationship between the CO_2 percentage of the respired air and the amount of air breathed (pulmonic ventilation).* That there is such a relationship has been most successfully demonstrated by R. W. Scott, who used for his purpose decerebrate cats.† The trachea was connected, through a T-tube provided with valves, with tubing leading to a large bottle and a Gad-Krogh spirometer, so that the animal breathed out of the bottle into the spirometer, these two being also connected together. The spirom-

*A distinction is somewhere drawn between pulmonic ventilation and alveolar ventilation, the former being the total amount of air that enters and leaves the lungs, and the latter, that which enters and leaves the alveoli. This distinction is based on the assumption that the capacity of the dead space may vary considerably from time to time, which, as pointed out elsewhere, is erroneous. For practical purposes pulmonic ventilation is the safer value to give.

†Decerebrate animals must be used in these experiments, since anesthetics markedly depress the activity of the respiratory center.

eter was made to record its movements on a drum, so that an accurate record of the depth and frequency of the respirations was secured. Samples of air were removed from the bottle by ground-glass plunger syringes at frequent intervals during the time that the animal was respiring into the tubing.

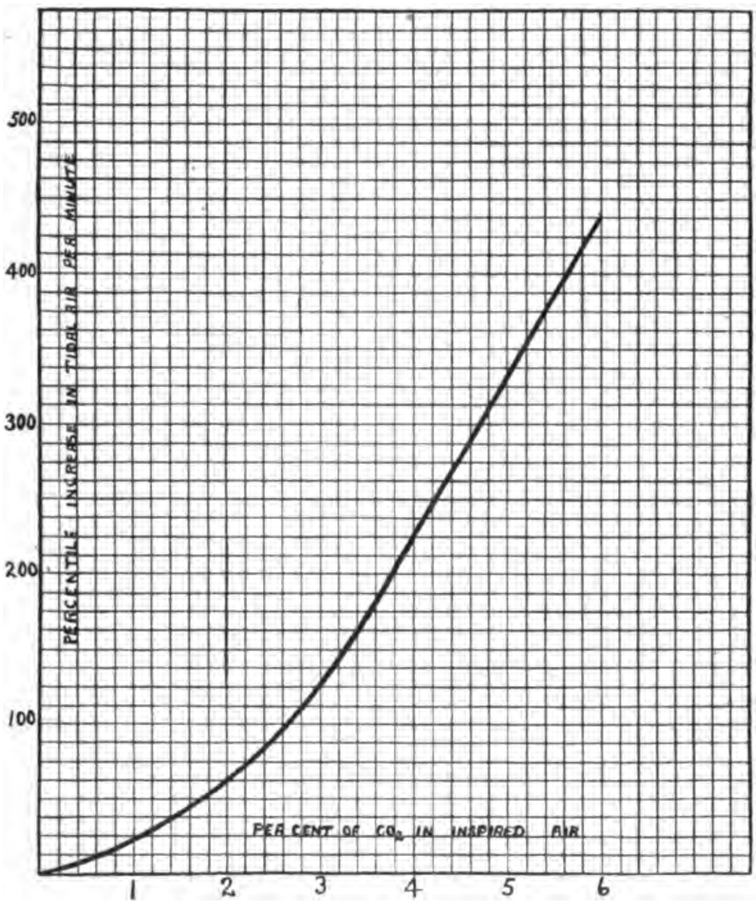


Fig. 129.—Composite curve obtained from the data on sixteen experiments, showing the respiratory response to CO_2 in the decerebrate cat. Abscissae = percentage of CO_2 in the inspired air. Ordinates = the percentile increase the tidal air per minute. (From R. W. Scott.)

The results are given in the accompanying curve (Fig. 129), which shows that there is a perfect correspondence between the CO_2 percentage in the air of the bottle and the pulmonary ventilation. Moreover, when the bottle was filled with O_2 instead of air to start with, the same results were obtained, showing that the CO_2 accumulation alone was responsible for the hyperpnea. In these cases the percentage of O_2 remaining in the

system after hyperpnea had become extreme, was far above that at which direct excitation of the center from O_2 deficiency is possible.

Experiments of a similar type had previously been performed by Porter and his pupils,[21] but their object was not so much to show the close parallelism between the CO_2 content of the respired air and the pulmonic ventilation as to demonstrate the changes produced in the sensitivity of the respiratory center in pneumonia.

Possibility that CO_2 Specifically Stimulates Center.—After showing that CO_2 acts as an excitant of the respiratory center, the question arises whether we are justified in the assumption that has been made tentatively that the action depends on the raising of the C_H of the blood, or whether it may be a specific action of the HCO_3 anion itself. Many attempts have been made to decide this question experimentally, the general principle of the experiments being to determine whether C_H of the blood runs parallel with the CO_2 content of the respired air and with the hyperpnea. Using the gas-chain method (page 31), Hasselbalch and Lundsgaard[22] found that the hyperpnea produced in rabbits by breathing in CO_2-rich air runs approximately parallel with the increase in the C_H of the blood, but on account of the experimental difficulties encountered they could not decide whether changes in C_H are *alone* responsible for the effect. These authors had previously demonstrated that changes in C_H can be induced in blood removed from the body by alterations in the CO_2 tension within the physiologic limits. An increase of one millimeter in CO_2 tension was found to cause an increase in C_H of 0.0065×10^{-7} (see page 27).

R. W. Scott's experiments, above referred to, have, however, yielded more definite results. By using the colorimetric method for determining C_H of the blood (see page 32), it could be readily shown, as is evident from the table (col. 8 in table), that a marked rise in C_H became evident when the inspired air contained 5 per cent or more of CO_2. That this rise was due to increase in the CO_2 tension was shown not only by finding a greater percentage of CO_2 (col. 15) in the blood, but also by being able to demonstrate that when CO_2-free air was bubbled through the blood removed during the dyspnea, C_H immediately returned to the normal, which it also did when the blood removed after the animal had breathed for a few minutes in outside air (col. 16). The CO_2 content likewise returned (col. 17). Had the increase in acidity been caused by nonvolatile acids—lactic, for example —these results, particularly the latter, could not have been obtained.

Although there is therefore no doubt that the C_H of the blood may be raised because of an increase in CO_2 in solution in the blood plasma—a CO_2 acidosis, as we may call it (see page 354)—this does not prove that the stimulation of the respiratory center is brought about solely by C_H.

The increase in the carbonate ion—HCO_3 ion—itself might also serve as a stimulus. That such is actually the case was demonstrated by finding that, if the C_H of the blood was first of all lowered by injecting alkali intravenously, hyperpnea still developed in proportion as the CO_2 accumulated in the inspired air; and that C_H of the blood, when the hyperpnea was at its highest, was *below* that of normal blood. Some other factor than C_H must obviously be responsible for this result. This must be the HCO_3 anion.

THE EFFECT OF REBREATHING CARBON DIOXIDE ON THE MINUTE VOLUME AND ON THE H-ION CONCENTRATION AND TOTAL CARBONATE CONTENT OF THE ARTERIAL BLOOD IN THE DECEREBRATE CAT

	PRELIMINARY PERIOD*								REBREATHING PERIOD						FIFTEEN MINUTES AFTER REBREATHING PERIOD	
Experiment number	Cat number	Weight in kg.	Time	Temperature per rectum	Respiration per minute	Tidal air per minute	P_H arterial blood	Total CO_2 content of arterial blood per 100 c.c.	Respiration per minute	Tidal air per minute	Percentile increase in tidal air per minute	CO_2 in inspired air	P_H arterial blood	Total CO_2 content of arterial blood per 100 c.c.	P_H arterial blood	Total CO_2 content of arterial blood per 100 c.c.
		kg.				*c.c.*		*c.c.*		*c.c.*	*per cent*	*per cent*		*c.c.*		*c.c.*
27	36	2.2	10.30 a.m.	38.5	48	864	7.45	30.3	64	3,968	540	6.45	7.25	53.4	7.4	34.2
28	33	2.2	1.50 p.m.	38.8	22	616	7.40	40.2	40	2,880	367	5.02	7.20	58.0	7.35	42.0
29	35	2.0	10.30 p.m.	38.6	28	784	7.40	38.1	40	3,840	389	5.35	7.25	51.1	7.4	40.0
30	32	2.3	11.00 p.m.	38.7	32	768	7.45	30.2	54	4,536	489	6.40	7.20	51.1	7.4	35.0
31	31	2.0	11.15 a.m.	38.8	32	896	7.40	32.0	60	4,800	436	5.95	7.20	50.0	7.4	35.2
		Average			32.4	785	7.42	34.1								

*P_H is the actual value given in the table. This is inversely proportional to C_H.

Further corroboration of the claim that the HCO_3 anion has a specific stimulating effect on the respiratory center that is independent of C_H, has been furnished by Hooker, Wilson and Connett.[23] These authors succeeded in keeping the centers of the medulla alive by perfusion with defibrinated blood through the blood vessels of the brain, and found that, although the respiratory movements of the diaphragm became depressed with a decrease and excited with an increase in C_H of the perfused fluid, a greater activity of the center was produced when this contained a high tension of CO_2 than with another fluid of the same C_H but with a low tension of CO_2. *We conclude that, although the C_H is the important respiratory hormone, the carbonate ion (HCO_3 anion) also has a stimulating influence.*

Relationship Among Acidosis Conditions, Alveolar CO_2 and Respiratory Activity.—It will be plain that variations in the respiratory hormone, whatever this may be, are associated with changes in the CO_2 content of the alveolar air. Closer examination has shown, however, that this relationship is by no means always so simple as in the instances just described, where increased respiration was found to be associated with an increase in alveolar CO_2. There are many cases where the reverse relationship obtains—namely, decreased alveolar CO_2 and hyperpnea. The whole question is very closely linked with that of the control of the reaction of the body fluids and with the etiologic factors in acidosis. When it is fully answered, many obscure clinical conditions in which respiratory disturbances occur will be much better understood than they are at present. On account of its great importance, considerable attention will be devoted in the next few pages to some of the researches which have been made bearing on the relationship between the CO_2 of the alveolar air and the various modified types of breathing that can be produced experimentally or become developed under altered physiologic conditions.

We shall consider these conditions in the following order: (1) Constancy of the alveolar CO_2 under normal conditions and during moderate variations in barometric pressure. (2) The quantitative relationship between an artificially induced increase in alveolar CO_2 tension (as by breathing CO_2-rich air) and the increased respiration. (3) The results of these observations will demonstrate a very precise relationship to exist between alveolar CO_2 tension and respiration, but if we proceed to repeat the latter observations under conditions where the accumulation of CO_2 in the inspired air is accompanied by oxygen deficiency (as by breathing in a confined space), we shall see that the relationship no longer holds, indicating that the oxygen deficiency has caused something to happen which disturbs it.

We shall find that the disturbing factor is accumulation of unoxidized acids in the blood, and this will naturally lead us to study the conditions in which such acids might develop; namely, (4) Breathing in rarefied air (mountain sickness). (5) Apnea. (6) Muscular exercise.

In succeeding chapters, we intend to review the work in these fields in considerable detail, partly because of its very important bearing on the general question of the control of the respiratory center and partly because of the light the observations throw on the nature of the mechanism involved in the adjustment of the C_H of the blood and tissues.

As we have seen, much work concerning the physicochemical principles involved in the control of the reaction of the blood has been contributed during recent years by physical and biological chemists, but much of this

work in our judgment fails to pay sufficient regard to the extraordinarily complicated conditions existing in the animal body, and more particularly, to correlate the purely physicochemical data with the numerous observations that have from time to time been recorded by physiologists regarding the behavior of the respiratory center. Physical chemists have recently, for example, gone so far as to define acidosis as a condition in which there is a diminution in the bicarbonate content of the blood induced by the discharge into it of fixed acids. This is going too far, for it fails to recognize acidosis due to an increase in the CO_2 of the blood. It is the molecular ratio $\left[\frac{H_2CO_3}{NaHCO_3}\right]$ which determines the tension of CO_2. When CO_2 is added to the blood, either experimentally by respiring the gas, or naturally owing to muscular exercise or to pathologic conditions in which there is a deficient excretion of CO_2, as in heart disease, the tendency of the equation to change, by increase of the numerator, is prevented partly by stimulation of the respiratory center, which gets rid of CO_2, and partly by an increase in the denominator. The respiratory center is so sensitive to slight increases in C_H that it becomes excited before a sufficient increase in H_2CO_3 has occurred to disturb the normal ratio $\left[\frac{H_2CO_3}{NaHCO_3.}\right]$. When fixed acids are added to the blood the denominator of the equation, $NaHCO_3$, is lowered and consequently C_H rises, and increased respiration occurs, lowering H_2CO_3 and thus reestablishing the ratio.

CHAPTER XLI

THE CONTROL OF RESPIRATION (Cont'd)

THE CONSTANCY OF THE ALVEOLAR CO_2 TENSION UNDER NORMAL CONDITIONS

Since a close relationship exists between the alveolar CO_2 tension and the respiratory activity, it is to be expected that the two would bear a strict proportionality to each other, and since the breathing under normal conditions does not vary much, the CO_2 tension should also be constant. Many observations show this to be the case. The tension is remarkably constant from day to day and even from month to month in the same individual, provided the physiologic conditions are the same. A slight seasonal variation is said to exist, a rise in the temperature of the environment of the individual causing a slight depression in the CO_2 tension, while a fall in temperature causes a slight rise (Haldane). These changes are independent of any demonstrable change in rectal temperature and, therefore, are probably due to the influence of the temperature on the skin.

Since it is the number of molecules of CO_2 in a given volume of alveolar air (i. e., the partial pressure or tension) that is of importance, it is only when the barometric pressure is the same that the *percentage* of CO_2 in the sample of alveolar air can be constant. To allow for this, all results are reduced to standard barometric pressure (760 mm. Hg). If the barometric pressure is lowered, there will have to be a higher percentage of CO_2 in the sample in order that there may be the same tension of this gas in the air of the alveoli; and *vice versa* when the barometric pressure is raised. The equation by which this tension, expressed in millimeters of mercury, is determined is: 100:760::a:p, where a is the percentage actually found in the air of the sampling tube and p the tension. A correction must also be introduced in this equation to allow for the vapor tension of the air in the alveoli, for of course H_2O molecules will behave like CO_2 molecules in causing a partial pressure.

When reduced to this standard, it has been found that the tension of CO_2 in the alveolar air remains constant under the different barometric conditions that obtain at the top of a mountain or at the foot of a deep mine. This is shown in the following table:

	(1) BAROMETRIC PRESSURE (MM. HG)	(2) CO_2 ACTUALLY FOUND IN DRY ALVEOLAR AIR (PER CENT)	(3) PARTIAL PRESSURE OF CO_2 IN MOIST ALVEOLAR AIR AFTER CALCULATING FOR BAROMETRIC PRESSURE
Top of Ben Nevis	646.5	6.62	5.23*
Oxford	755	5.95	5.53
Foot of Dolcoath Mine	832	5.29	5.48
Compressed air cabinet	1260	3.52	5.64

*The figures in this column are arrived at by the formula: $\frac{B' - A \times P'}{B} = P$, when P = figures in last column; B' = figures in first column; A = aqueous tension of alveolar air; P' = figures of second column; B = barometric pressure at sea level. A is obtained from tables giving the aqueous tension at different temperatures.

Changes in the frequency of breathing that are within physiologic limits have no influence on the tension of alveolar CO_2, provided that exactly the same time is taken in performing the forced expirations during which the samples of alveolar air for analysis are removed.

The Degree of Sensitivity of the Respiratory Center to Changes in the CO_2 Tension of the Alveolar Air

This can be determined by observing the alterations produced in the volume of air that actually enters the alveoli (alveolar ventilation) during breathing in atmospheres containing different percentages of CO_2. In man an increase of from 0.2 to 0.3 per cent in the alveolar CO_2 is sufficient to double approximately the alveolar ventilation; or, more precisely, an increase of ten liters in the air entering and leaving the alveoli per minute is caused by raising the alveolar CO_2 tension by from 2.2 to 3.1 mm. Hg (Douglas, etc.)[24].

THE ALVEOLAR CO_2 TENSION DURING BREATHING IN A CONFINED SPACE

We have already employed similar experiments in ascertaining whether CO_2 accumulation or O_2 depletion is responsible for the hyperpnea produced under these conditions. We concluded for the former, but now on closer examination we shall see that, although our conclusion was correct, the deficiency in O_2 also has an indirect effect on the respiratory center. This is revealed by the fact that *the tension of the CO_2 in the alveolar air does not increase in proportion to the observed increase in pulmonary ventilation.* We must conclude that the decrease in O_2 has some effect. How may this be explained? Two possibilities exist: (1) that the O_2 want has caused organic acids to accumulate in the blood and so raise the C_H; and (2) that in the absence of a certain tension of

O_2 the excitability of the center is raised (i. e., its "threshold" lowered), so that it becomes stimulated by C_H, to which ordinarily it does not respond. We shall now proceed to examine the experimental evidence bearing on these possibilities.

By examination of the alveolar air of an individual confined in a pneumatic cabinet in which the barometric pressure is gradually lowered, it has been found that although the CO_2 tension remains constant for a considerable range (cf. page 356), it begins to fall when the barometric pressure has reached about 550 mm. Hg. At this pressure the tension of O_2 in the alveolar air will be 62 mm. instead of its normal of about 105 at atmospheric pressure. Below it the alveolar CO_2 tension quickly falls, and at the same time hyperpnea becomes evident, although the person himself may be unaware that he is breathing more deeply. If this experiment is repeated with the difference that, as the pressure is lowered, an excess of O_2 is introduced into the chamber, the hyperpnea does not supervene until a barometric pressure has been reached that is distinctly lower than when no excess of O_2 is present, and at the same time the CO_2 tension in the alveolar air remains unchanged. The explanation of this result is that by lowering the O_2 tension in the alveolar air and, therefore, in the blood and tissues, oxidative processes become depressed so that unoxidized acids, such as lactic, accumulate in the blood and by adding their effect to that of the CO_2 serve to raise the C_H of the blood. As a result, the respiratory center becomes excited, hyperpnea supervenes, and the volatile CO_2 is removed from the blood into the alveolar air. On supplying O_2 artificially, this failure of proper oxidation does not set in and breathing goes on normally.

There should be a stage in the above experiment during which the CO_2 tension of the alveolar air is *increased*—namely, when the fixed acids first appear and decompose the carbonates of the blood. This stage has, however, not been detected. When a person is kept in such a chamber for some time at a pressure which causes a diminution in the alveolar CO_2 tension, the tension does not immediately return to its normal level when atmospheric air is again breathed, indicating that the fixed acids are only slowly got rid of.

The second hypothesis—namely, that the O_2 deficiency directly raises the excitability of the respiratory center—has many advocates, among them Lindhard,[25] who found that, when the percentage of O_2 in the alveolar air was raised, a higher percentage of CO_2 was necessary to cause an increase in the ventilation of the lungs, and conversely, that a distinct increase in the excitability of the center occurred when the inspired air contained less than the normal percentage of O_2. Although it is admitted by Haldane and his school that such alteration in the excitability

of the respiratory center to the C_H of the blood may occur after long-continued exposure of the center to the changed tension of O_2, yet they deny that such alteration can occur as a temporary condition. These workers found that, in order to raise the pulmonic ventilation by 100 per cent, the increase in the alveolar CO_2 tension required was practically the same (0.3 per cent) when the inspired air contained 20 per cent of O_2 as when it contained 54 per cent.

In the observations already referred to on the decerebrate cat, R. W. Scott[20] has secured some evidence that would seem to support Haldane's contention. He found that the response of the respiratory center to the percentage of CO_2 in the respired air was exactly the same whether the latter contained a low (13-14) or a high (30 and over) percentage of O_2. The possibility that the excitability of the respiratory center is affected directly by the O_2 tension is to be considered as one of the most important problems awaiting solution.

Even if it may have a certain influence on the excitability of the respiratory center, O_2 deficiency *per se* can serve as a direct stimulus of the center only when it is of extreme degree. Much light has been thrown on the relationship of O_2 to respiratory activity by observing the respirations during breathing in and out of rubber bags through soda lime absorption bottles of sufficient size to remove the CO_2. We have already seen that even the general results of such observations (page 349) show clearly how much more potent a respiratory stimulant is accumulation of CO_2 than deficiency of O_2. More particular investigation in which the alveolar air is analyzed bears out these conclusions and at the same time indicates the exact conditions under which organic acids become developed.

With a very small bag (a few liters' capacity) hyperpnea of a distressing type but without cyanosis supervenes in a few minutes, and the alveolar air contains perhaps as low as 6 per cent O_2 and 4 per cent CO_2. Of still greater interest and significance, however, is the fact that the ratio between the volume of CO_2 excreted and of O_2 absorbed (respiratory quotient) during the hyperpnea is raised considerably above unity, indicating that an excessive excretion of CO_2 must be occurring. This result is explained by assuming that the deprivation of O_2 causes large quantities of fixed acids to be produced, and that these expel CO_2 from the blood more quickly than the O_2 is absorbed. In corroboration of this explanation, it has been observed that, after outside air is breathed for some time following the above experiment, the respiratory quotient becomes very *low*, so that CO_2 must now be accumulating in the blood.

If the above experiment is repeated with a *larger* bag (about 200 liters), so that the O_2 falls slowly, the breathing can be maintained for

a much longer period *without any evident symptoms of hyperpnea,* even though the O_2 percentage in the alveolar air may fall as low as in the previous experiment, and there are marked symptoms of O_2 want, such as cyanosis, twitching of the muscles of the hands, lips, etc. The respiratory quotient does not become abnormal in this experiment indicating that no expulsion of CO_2 from the blood can have occurred as in the previous experiment. The cause for the virtual absence of hyperpnea in this experiment is no doubt that the more gradual reduction in O_2 of the alveolar air and therefore of the blood did not bring about the accumulation of lactic acid at a rate that was greater than that at which the CO_2 was got rid of into the alveolar air.

BREATHING IN RAREFIED AIR; MOUNTAIN SICKNESS

In considering the part played by fixed organic acid in the control of the C_H of the blood, the most important results have been secured by observations on the condition of individuals living at high altitudes. As is well known, under these conditions certain symptoms are likely to develop, the condition being known as *mountain sickness.* The great interest which physiologists have taken in this subject has been owing, not so much to the importance of the observations in connection with the condition itself, as to the light which they throw on the mechanism of respiratory control and on the cause for abnormal types of breathing.

More or less hyperpnea, especially on exertion, soon appears in a rarefied atmosphere, and the alveolar CO_2 tension assumes a value considerably below the normal. For example, at sea level the minute volume of air breathed in one individual was 10.4 liters, and the alveolar CO_2 tension 39.6 mm. Hg. After being some time on Pike's Peak, where the barometer registers only 459 mm. Hg, Douglas[26] found the minute volume of air to be 14.9 liters, and the alveolar CO_2 tension 27.1 mm. Hg. At first sight the above statement may seem to contradict one previously made, to the effect that the alveolar CO_2 tension remains constant at different barometric pressures. This applies, however, to the immediate effects, whereas we are now considering the later effects. The important point is: How are we to reconcile with the above hypothesis the fact that a diminution in the alveolar CO_2 tension should be accompanied by hyperpnea? A solution of the seeming contradiction will not only be of importance in connection with our present problem, but will assist us in the investigation of the clinical conditions of hyperpnea, in which likewise a diminished CO_2 alveolar tension is often observed. ***Mountain sickness may indeed be considered as an intermediate condition between the physiologic and the pathologic.***

From what we have learned we should expect the above result to be

dependent upon an increase in the nonvolatile acid content of the blood That such is really the case has been conclusively shown both by titration methods and by observing the dissociation curve of hemoglobin, which, as will be explained later (see page 386), may be made to serve

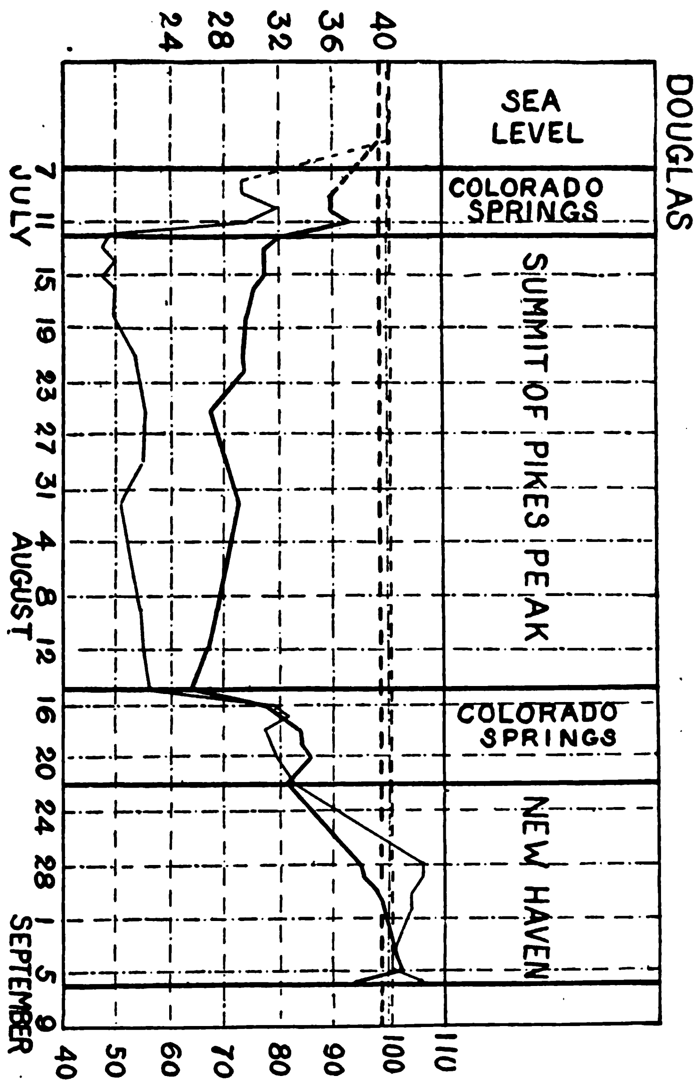


Fig. 130.--The horizontal interrupted lines represent the mean normal alveolar CO_2 and O_2 pressures at sea level (i. e., Oxford and New Haven); the thick line, alveolar CO_2 pressure; and the thin line, alveolar O_2 pressure. (From Douglas, Haldane, Henderson, and Schneider.)

as an index of the H-ion concentration of the blood. The exact chemical nature of the nonvolatile acids that accumulate in the blood is not as yet known. Two types of acid can be thought of, either unoxidized organic

acids, of which lactic acid may be taken as the representative, or inorganic substances, like the acid phosphates. That it is not lactic acid is shown by both direct and indirect evidence. The direct evidence has been furnished by Ryffel, who was unable to find any increased percentage of this substance either in the urine or in the blood of persons who had been living for some time in the famous Regina Margherita hut on Monte Rosa.[27] The indirect evidence has been furnished by observing the time that it takes after the individual has started breathing the rarefied air for the alveolar CO_2 tension to fall, as well as that required to bring about the recovery to the normal when he descends to sea level. The following curve, which is self-explanatory, will illustrate these points.

Thus, on Pike's Peak, where the barometric pressure is 459 mm. Hg, the CO_2 tension after an initial fall took about seven days before it came to its permanent level for that barometric pressure, and fourteen days elapsed after descending from the mountain before the sea-level tension had been regained. The slow nature of these changes, when compared with the rapid changes observed in the experiment with the bags already alluded to (page 358), *shows clearly that lactic acid can not be responsible for the increase in H-ion concentration in mountain sickness.* By exclusion it would appear that the increase in C_H is the result of an excess of fixed inorganic acid (H_3PO_4) in the blood dependent on a disproportionate excretion of bases by the kidneys during the period of acclimatization to the rarefied air.

Other observers aver that the acidosis does not really exist, but that the excitability of the respiratory center itself becomes raised (its threshold lowered), so that it responds more readily to the normal C_H of the blood. It has been stated that the increase in excitability of the center is dependent upon the action of the intense light rays at high altitudes—the erythema of the skin, etc., being evidence of this exciting action of light. The constant irritation of the skin, these authors say, serves by stimulation of afferent nerves to maintain a hyperexcitability of the respiratory center. Others believe that the hyperexcitability of the center is a direct result of the maintained O_2 deficiency. The balance of evidence, however, stands in favor of the view that the phenomena of mountain sickness depend on changes occurring in the inorganic nonvolatile acids of the blood. The other phenomena of this interesting condition will be discussed elsewhere (page 399).

APNEA

If a man breathes forcibly and quickly for about two minutes, he will experience no desire to breathe for a further period of about the

same duration—he becomes apneic. When the desire to breathe returns, the breathing is at first very shallow, but gradually becomes more marked, until at last normal respiration is reestablished. If a sample of alveolar air is removed at the time when the desire to breathe returns, it will be found to contain a very small percentage of O_2 indicating that for some time previous to the onset of breathing there had been in the alveolar air, and therefore in the blood, so low a percentage of O_2 that if O_2 deficiency could stimulate breathing, this would have started much earlier than it actually did. A curve showing the results of such an experiment by Haldane is given in Fig. 131. The person may begin to show symptoms of O_2 want, such as cyanosis, before the desire to breathe returns, which furnishes strong proof that O_2 want itself can not serve as a stimulus to the respiratory center. The failure of the center to act must rather be due to the lowering of the C_H consequent upon the removal of CO_2 from the blood by the forced respiration which preceded the apnea—*washing out of the* CO_2, as it is called. That this has really occurred can readily be shown by estimating the CO_2 content of a sample of alveolar air collected by having the subject make a forced expiration early in apnea. Extremely low values along with a respiratory quotient (page 547) of about 0.2 are often found, whereas, during the preceding forced breathing while the CO_2 is being washed out, the quotient is often ten times as great—viz., 2.0.

As would be expected, the low O_2 percentage present in the alveolar air toward the end of the apneic pause is not without some effect, indirect though it may be, on the excitability of the respiratory center. This accounts for the fact that the alveolar air, at the moment the desire to breathe returns, usually contains a *lower* percentage of CO_2 than the normal, indicating that some nonvolatile acid must have accumulated in the organism so as to raise the C_H of the blood, and thus require a lower tension of CO_2 to overstep the threshold of excitability of the respiratory center. In agreement with this explanation it has been found that, if the last two or three forced respirations preceding the apnea are made in an atmosphere of O_2 instead of air, so as to fill the alveoli with O_2, the apnea can be maintained for a very much longer period; and when the natural desire to breathe returns, the CO_2 tension of the alveolar air, instead of being below the normal, is above it. The effect of O_2 in prolonging apnea must, therefore, be dependent on the fact that it prevents the accumulation in the organism of the unoxidized acids, leaving to CO_2 alone the function of raising the C_H in the blood to the level required to excite the respiratory center. By this means the period during which the breath can be held after breathing O_2 is sometimes

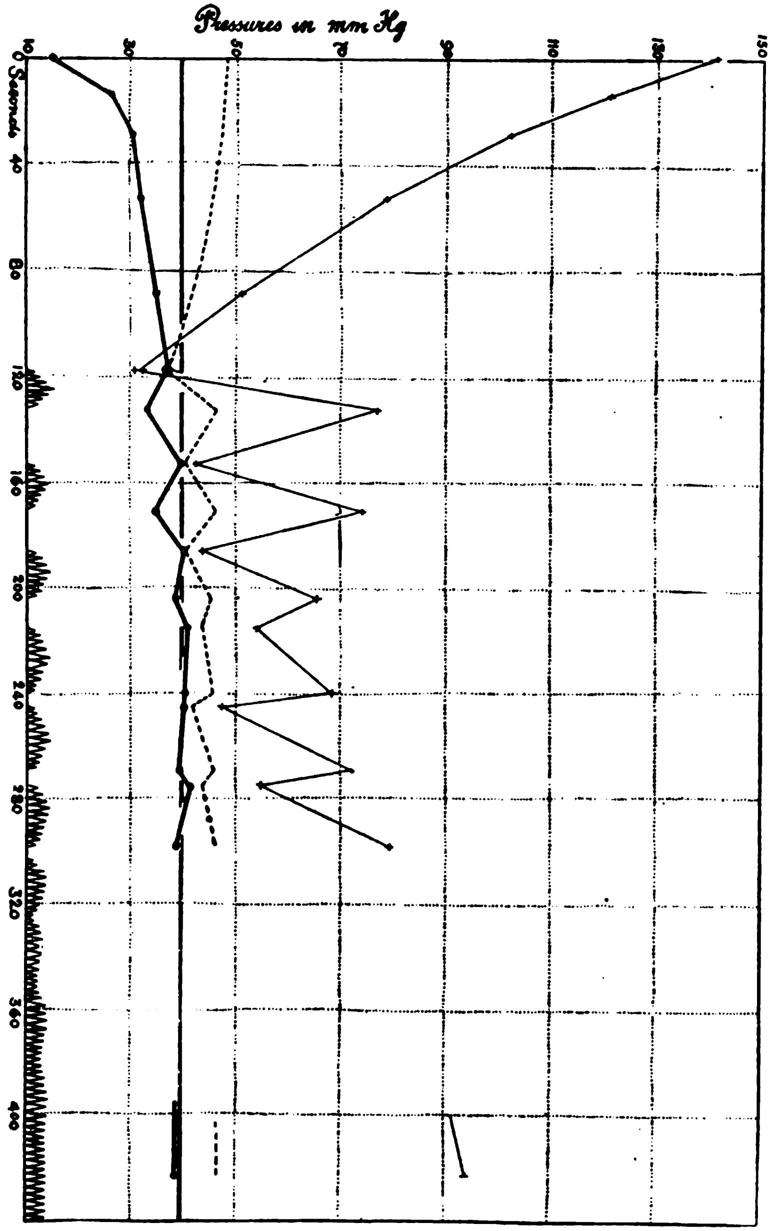


Fig. 131.—Curves showing variations in alveolar gas tensions after forced breathing for two minutes. Thin line = O_2 tension; thick line = CO_2 tension. Double line = normal alveolar CO_2 tension. Dotted line shows the alveolar CO_2 tension at which breathing would recommence at the end of apnea with the alveolar O_2 pressures shown by the thin line. The actual breathing is indicated at the lower part of the figure. It is periodic to start with. (From Douglas and Haldane.)

phenomenal; in one individual, for example, after breathing forcibly for a few minutes and then filling the lungs with O_2, apnea lasted for eight minutes and seventeen seconds.

The Supposed Nervous Element in Apnea

It is necessary to point out that, prior to the elaboration of accurate methods for the investigation of the chemistry of respiration, many physiologists interpreted the apnea following forced breathing as the result of a sort of inhibition of the respiratory center brought about by its repeated stimulation by afferent nervous impulses transmitted to it along the vagus nerves, these impulses being set up by the frequent collapse and distention of the alveoli acting on the terminations of the nerve. In justification of the nervous interpretation of apnea, it was claimed by the earlier observers that it could not readily be produced in animals after severing both vagus nerves. More recent work has shown that this is not an accurate observation, for if the severing of the vagi is accomplished not by cutting but by freezing, then apnea is as readily produced as in an intact animal (Milroy).[28]

That chemical and not nervous factors cause the apnea is further demonstrated by the well-known experiment of Fredericq, who, after ligating the vertebral and one of the carotid arteries in two dogs, anastomosed the central end of the remaining carotid of the one to the peripheral end of the carotid of the other animal, thus establishing a crossed circulation. He then found that by applying forced artificial respiration to the one animal, the apnea which supervened affected the other animal and not that to which the artificial respiration had actually been applied. Another proof of the chemical theory of apnea is furnished by the observation that if forced breathing is performed in an atmosphere containing CO_2 in about the same partial pressure as in the alveolar air, no apnea supervenes, and if the experiment is repeated several times with progressively declining percentage of CO_2 in the air each time, the length of the apneic pause proportionally increases as the CO_2 pressure in the inspired air diminishes.

Although in the foregoing account we have adopted Haldane's view that oxygen deficiency *per se* can act as an excitant of the respiratory center only when it is of extreme degree, it should nevertheless be pointed out that studies by A. S. Loevenhart on the action of cyanides on the respiratory center have led him to conclude that interference with oxidative processes may be a more potent factor in its stimulation than the experiments in which oxygen-poor atmospheres are respired would lead us to expect. We must await further evidence before a final verdict is pronounced on this most perplexing problem of modern physiology.

CHAPTER XLII

THE CONTROL OF RESPIRATION (Cont'd)

THE EFFECT OF MUSCULAR EXERCISE ON THE RESPIRATION

During muscular exercise the pulmonic ventilation increases to an extraordinary extent. At rest an average man respires 6 to 8 liters of air per minute, but during walking on the level at the rate of 5 kilometers an hour, this figure may increase to about 20 liters.

The first investigations as to the cause of the relationship between muscular activity and pulmonic ventilation were made by animal experiments in which tetanus of the muscles of the hind limbs was produced by electric stimulation of the spinal cord. The problem was to find out what serves as the means of correlation (nerve reflex or hormone control) between the muscular activity and the respiratory activity. By cutting the spinal cord above the point of stimulation, it was found that the tetanus was still accompanied by as marked a hyperpnea as before. On the other hand, when the spinal cord was left intact but the blood vessels of the limb were ligated, no hyperpnea followed the tetanus. Evidently therefore the pathway of communication is the blood.

The next step was to seek in the blood for the substance or hormone that acted as the respiratory excitant, and naturally the first possibility considered was a change in the gases of the blood, either a deficiency of O_2 or an increase in CO_2. Direct examination of the blood for the quantity of these gases, however, yielded results which were quite contrary to such an hypothesis. It was found that the percentage of O_2, if anything, was slightly increased, and that of the CO_2, if anything, diminished. Moreover, when the expired air was analyzed during the hyperpnea, the *percentage* of CO_2 contained in it was distinctly below the normal average, and the percentage of O_2 above it. Evidently, therefore, the *amount* of gases in the blood has nothing to do with the excitation of the respiratory center, and the conclusion drawn by the earlier investigators was to the effect that the exciting substance carried from the active muscles to the respiratory center must be some unusual metabolic product, possibly the lactic acid produced by contraction.

It was further found, by examination of the respiratory quotient, that

an excess of CO_2 was being expired during the work and immediately after it, but that this was subsequently followed by a much lower quotient, indicating that CO_2 was being retained. Such a result would be in conformity with the view that an acid such as lactic is discharged into the blood, on the carbonates of which it would act as explained on page 355. Breathing in and out of a small rubber bag causes the same alterations in the respiratory quotient (see page 358).

That lactic acid is actually produced by contracting muscle could not, however, be shown by all investigators, and it was not until some years later that Fletcher and Hopkins[29] clearly demonstrated the conditions under which it may appear in active isolated muscle. These observers found that lactic acid is produced in excised muscles only when the muscular contraction occurs in a deficiency of O_2. When it occurs in an adequate supply of O_2, CO_2 instead of lactic acid is produced.

Taking these facts together with what we already know concerning the conditions under which the respiratory center reacts to conditions which presumably cause a change in the C_H of the blood, we may formulate the hypothesis that respiratory activity during muscular exercise is due to a slight increase in the C_H of the blood, and that this increase is owing *partly to an actual increase in CO_2 production by the acting muscles and partly to the production of lactic acid.* Such an hypothesis would satisfactorily explain why the actual amount of CO_2 in the blood might be below the normal during muscular exercise, for the CO_2 would be "washed out" from the blood by the hyperpnea induced by the increase in C_H.

The obvious method of putting this hypothesis to the test is to examine the alveolar CO_2 tension and the respiratory quotient under various conditions of muscular activity. The results of such observations are given in the accompanying table.

	(1) O_2 used in c.c. per min.	(2) CO_2 produced in c.c. per min.	(3) R. Q. $\frac{\text{vol. } CO_2}{\text{vol. } O_2}$	(4) CO_2 in alveolar air	(5) Total alveolar ventilation in liters per min.
1. During rest, standing	328	264	0.804	5.70	5.80
2. Walking at the rate of 3 kilometers per hour	780	662	0.849	6.04	13.6
3. Walking at the rate of 5 kilometers per hour	1065	922	0.866	6.10	18.8
4. Walking at the rate of 6 kilometers per hour	1595	1398	0.876	6.36	27.6
5. Walking at the rate of 7 kilometers per hour	2005	1788	0.891	6.20	35.6
6. Walking at the rate of 8 kilometers per hour	2543	2386	0.938	6.10	48.2

In the first column is given the O_2 used in c.c. per minute. Among other things these figures indicate the actual amount of work done. In the second column is given the CO_2 production in c.c. per minute. By dividing the figures of the second column by those of the first, we obtain the figures of the third column, representing the respiratory quotient. The fourth column gives the CO_2 content of the alveolar air, and the last column the total *alveolar* ventilation in liters per minute.

Taking for the present the figures in the first and fourth columns and postponing a consideration of the respiratory quotient, it will be noted that, as the muscular work increases up to a total consumption of about 1600 c.c. of O_2 per minute, the CO_2 percentage in the alveolar air steadily increases. The question arises, does the alveolar ventilation increase in proportion to the increase in CO_2 tension? If it does so, increase in CO_2 tension in the blood can be held solely responsible for the hyperpnea (i. e., a pure CO_2 acidosis); whereas if the hyperpnea is greater than can be accounted for by the increase in CO_2 tension, other acids must be partly responsible for the acidosis. By making this same individual breathe atmospheres containing different percentages of CO_2 it was found that to produce a doubling of the alveolar ventilation it required an increase amounting to 0.33 per cent of an atmosphere of CO_2 in the alveolar air (see also page 357). When we examine the above figures during muscular exercise, however, we find that a rise in alveolar CO_2 from 5.70 to 6.36 (i. e., 0.66 per cent) multiplied the normal alveolar ventilation by considerably more than four times, whereas had it been entirely due to an increase in CO_2, it should not have been more than twice as much. Evidently therefore, some other factor than CO_2 tension must have been responsible for the increased respiratory activity. This conclusion is further confirmed by examination of the alveolar CO_2 during very strenuous muscular effort, when a relative *decrease* in the CO_2 percentage becomes apparent.

If it is true that the exciting agency has been dependent partly on an increase in the CO_2 tension of the blood, and partly on the production of nonvolatile organic acids (lactic acid), we should expect that immediately after discontinuing the muscular exercise the CO_2 tension of the alveolar air would fall to a level distinctly below normal, that it would only slowly recover thereafter, and that further exercise before the recovery had occurred would produce only a slight increase in alveolar CO_2. These results we should expect because of the much slower rate at which the nonvolatile organic acid is got rid of from the organism, compared with the volatile CO_2. By actual experiment these suppositions have been found to be correct, as is shown in the following table.

	TIME AFTER DISCONTINUING A BRIEF PERIOD OF MUSCULAR EXERCISE	ALVEOLAR CO_2 TENSION IN MM. HG
1st Period:	10″	49.2
	3′ 0″	35.4
	6′ 30″	35.3
	12′ 30″	35.8
2nd Period:	10″	38.9
	3′ 0″	33.7
	6′ 30″	34.4
3rd Period:	10″	36.9
	3′ 0″	34.4
	8′ 30″	32.4
	18′ 30″	33.7
	24′ 0″	36.2
Normal resting:		39.0

(Douglas.)

In this table the figures of Period 1 represent the alveolar CO_2 tension in mm. Hg immediately following a period of strenuous work. The figures in Period 2 are for the same individual again performing the same amount of work with, however, only a short period of rest intervening, and the figures of the third period are a repetition of the same conditions. It will be observed that the muscular exercise at first raised the alveolar tension of CO_2 from the normal of 39 mm. to 49.2 mm., but that in three minutes after the work had been discontinued the tension was considerably below the normal. During the second period of muscular exercise the CO_2 in the alveolar air collected immediately after the effort did not increase above the normal level, and in the third period the increase was still less—results which are entirely in conformity with the view that as a consequence of the first period of muscular exercise nonvolatile organic acids had accumulated in the blood, so that to produce the required respiratory activity in the second and third periods a much less increase in CO_2 tension was required.

We may sum up the conclusions which these observations justify by stating that during muscular exercise the C_H of the blood becomes slightly increased because of the liberation into it of CO_2 and of lactic acid from the acting muscles. The respiratory center is, however, so sensitive to the slightest increase in C_H that it immediately responds and produces hyperpnea, with the result that the volatile CO_2 is so washed out of the blood that the C_H is held down in spite of the continued production of acid substances by the muscles. The more strenuous the exercise, the less able is the O_2 content of the blood to keep pace with the metabolic activity of the muscles, so that relatively more and more lactic acid is produced, necessitating therefore a greater and greater washing out of CO_2.

The readiness with which CO_2 can be got rid of prevents the hormone which excites the respiratory activity from continuing to act after it is no longer required. Provision for the removal of a hormone after its activity has been displayed is of course essential to efficient correlation of function, and is seen in the case of other hormones, such as epinephrine and secretin, whose discontinuance of action is effected by their destruction in the blood (see page 745).

Direct evidence that lactic acid is formed during strenuous muscular exercise in man has been furnished by Ryffel.[30] Blood removed from a person immediatey after running at full speed for about three minutes contained 70.8 milligrams of lactic acid per 100 c.c. of blood, the normal amount being 12.5 milligrams. Much of the lactic acid accumulating in the blood is no doubt got rid of by oxidation, but a large part of it is also excreted by the urine, in which it was found by Ryffel in considerable amount after strenuous muscular exertion.

Finally, let us consider for a moment *the behavior of the respiratory quotient.* This ratio rises early in the muscle work (Table on page 367), indicating that more CO_2 is being excreted than O_2 absorbed. After the work is discontinued, it usually falls below the normal because of retention of CO_2 to take the place of the lactic acid that is being gradually used up or excreted. A similar fall may sometimes occur in the respiratory quotient *during* muscular exercise, if this is continued for a long time. It probably indicates that a balance has been struck between the production of lactic acid in the muscles and the loss of this substance by oxidation. In any case it is a significant occurrence, for it coincides with the great improvement in the subjective sensations accompanying muscular exercise. It occurs, for example, at the same time as the appearance of the "second wind," when the circulatory and respiratory distress experienced during the earlier stages of strenuous muscular exertion disappear. The stages prior to the second wind correspond to the period when considerable quantities of free CO_2 are being got rid of from the blood and are probably creating a temporary maladjustment of the C_H which acts on the various medullary centers. If by forced breathing much of this CO_2 is discharged before the muscular exercise is undertaken, the initial hyperpnea is not nearly so marked.

CHAPTER XLIII

THE CONTROL OF RESPIRATION (Cont'd)

PERIODIC BREATHING

Types of Periodic Breathing

In the best known of these, called Cheyne-Stokes respiration, a period of hyperpnea supervenes upon one of apnea, each period following in regular sequence. After an apneic period, the breathing begins at first faintly, gradually becomes more pronounced until it is markedly exaggerated, and then fades off again to the apneic pause. Sometimes the apneic period is immediately followed by one of intense hyperpnea, there being no gradual increase in the respiratory movements. Between these two types all varieties of the condition are met (Fig. 132).

The conditions in which periodic breathing occurs may be divided into physiologic and pathologic groups. Of the physiologic conditions the following may be taken as examples: (1) Breathing in an atmosphere containing a deficiency of O_2; thus, periodic breathing is very readily produced in persons living in rarefied air. (2) The initial breathing following an apnea induced by forced ventilation of the lungs. In this postapneic periodicity, the apneic periods may at first be quite marked, but as breathing returns they become gradually shorter and the breathing intervals gradually longer, until normal respiration is restored (Fig. 131). (3) Breathing through a long tube having a small vessel containing soda lime inserted between the tube and the mouth, the whole capacity of this vessel and tubing being about a liter. This will cause periodic breathing in persons that are susceptible to oxygen deficiency. Even breathing through the tube without soda lime will sometimes cause a periodic type of breathing in such individuals.

The *pathologic* conditions in which periodic breathing becomes developed are particularly those associated with renal disease and cerebral hemorrhage. In many of these cases, the periodic breathing does not appear to depend on the same factors as are concerned in the experimental types. The symptoms would rather appear to depend on some influence of the higher cerebral (supranuclear) centers on the respiratory center. At least some other evidence of disturbance of the cerebral functions is always forthcoming, such as a slight paralytic shock, and the

periodic breathing is nearly always aggravated during sleep. Many of these cases are greatly benefited by administration of caffeine.

In both the physiologic and the pathologic groups, the breathing may develop a periodic character only when the person is asleep, and even normal people, particularly infants or very old people, may exhibit it to a certain degree.

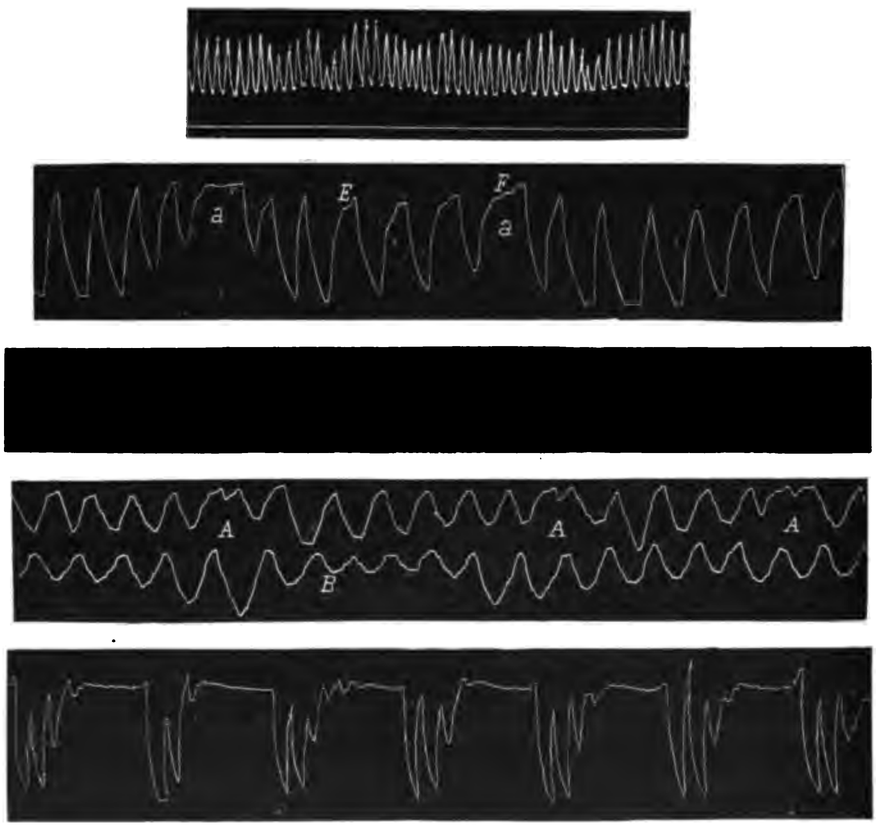


Fig. 132.—Various types of periodic breathing. (From Mosso's "Life of Man in the High Alps.")

Causes of Periodic Breathing

Great interest attaches to an investigation of the causes of periodic breathing, but it can not be claimed that any perfectly satisfactory explanation has as yet been offered. Pembrey[31] attributes it to a diminished excitability (a raised threshold) of the respiratory center due to faulty blood supply, the supposition being that, when thus suppressed, the normal C_H of the blood is unable to excite the center, so that breathing stops. During the resulting apnea, CO_2 again accumulates until it has

raised the C_H sufficiently to excite the depressed center. Hyperpnea follows, causing a washing out of the CO_2 and a resulting diminution of the effective stimulus, so that again the center fails to be stimulated and apnea supervenes, and so on. Support for this explanation would appear to be furnished by the fact that, when patients exhibiting periodic breathing are made to breathe an atmosphere containing a high percentage of CO_2, the periodicity of the breathing may give place to regular breathing; a result which may also be obtained by making such patients breathe in atmospheres rich in oxygen. In the former case, the stimulus is raised to meet the depressed excitability of the center; in the latter, the excitability of the center is increased because of better blood supply so that it is enabled to react to the diminished stimulus. But even granted that the excitability of the center is depressed, it is difficult to see why this should occasion a periodic type of breathing unless we assume that it is only when stimulus (i. e., C_H of blood) and threshold of excitability of the center are adjusted at a certain physiologic level that smooth and continuous action can go on.

Haldane and his school aver that there is no permanent alteration in the excitability of the center, but that the periodicity is due to several causes, which do not always operate to the same degree in the different conditions in which such periodicity exists. To study these causes the exact conditions existing in the various types of periodic breathing that can be produced experimentally in man have been investigated.

The most simple to consider first is the periodic breathing that is produced in a person susceptible to O_2 want, by breathing through a tube and bottle (of a total capacity of 1 liter), containing soda lime. In such a case no outside air enters the lungs, for what we have really done, besides providing for the absorption of CO_2, is greatly to prolong the dead space. The oxygen tension of the rebreathed air, therefore, quickly falls, until at last a point is reached at which the respiratory center is directly stimulated by O_2 deprivation, as we have seen it to be when this falls to a sufficiently low level (see page 350). The deep breaths (hyperpnea) which follow, being of greater volume than 1000 c.c., cause outside air to be inspired so that the O_2 want is made good and the hyperpnea again disappears, possibly to the extent of apnea, for now, in consequence of a coincident "washing out" of CO_2, there has been a lowering of the C_H of the blood below the threshold value. During the apnea the O_2 is rapidly used up, till a point is reached at which the center again becomes excited. In such an experiment the effect of O_2 want becomes very marked, as shown by the intense cyanosis which develops.

That breathing under these conditions should be periodic and not

merely show a steadily increasing hyperpnea is probably due to the unequal rates at which the O_2 and CO_2 tensions change in the blood. Because of a "buffer action" the latter fluctuates much less than the former. Another cause for the periodicity is no doubt the delay between the gas exchange in the lungs and the arrival of the blood in the brain. When the O_2 tension of the blood supplying the respiratory center falls to so low a level that excitation of the center occurs, the resulting increased breathing aspirates outside O_2 into the alveoli. After a moment or so, the O_2 is carried by the blood to the center, so that its stimulation by O_2 deficiency is removed, and it is left in a condition in which it fails to discharge any impulses, since there is a subnormal C_H of the blood as a consequence of the lowering of the CO_2 tension produced by the hyperpnea. A little time must now elapse before the CO_2 again rises or the O_2 falls sufficiently to excite the center.

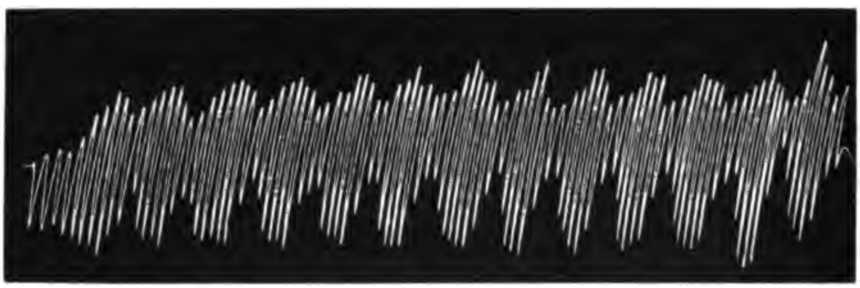


Fig. 133.—Quantitative record of breathing air through a tube 260 cm. long and 2 cm. in diameter. (From Douglas and Haldane.)

A similar although less marked degree of periodic breathing can sometimes be obtained by merely respiring through a long tube without any provision for the absorption of CO_2. In this case it is more difficult to explain the cause of the periodic breathing, but that the main factor concerned is one of O_2 deprivation is evidenced by the fact that in this as in the previous experiment, the periodic nature of the respiration is immediately changed to the regular breathing if O_2 is introduced into the tube. The interest of the experiment lies in the fact that a similar relative elongation of the dead space is probably accountable for the periodic breathing seen in the winter sleep of hibernating animals. During this condition, on account of the depression of metabolism less O_2 is required and less CO_2 is produced, so that the exchange of gases through the pulmonary endothelium is greatly diminished. The dead space, however, remains of the same capacity, which amounts to the same thing as if the latter had been prolonged under unchanged conditions of pulmonary gas exchange.

The explanation for other types of experimental periodic breathing is much less satisfactory. Important evidence that changes occurring in the tensions of O_2 and CO_2 in the alveolar air and therefore in the arterial blood of the respiratory center are largely responsible for periodic breathing has been secured by studying the condition that develops after a period of apnea produced by voluntary forced breathing. The results of such observations are given in the curve shown in Fig. 131.

The thin line represents the O_2 tension of the alveolar air, the thick line the CO_2 tension. The double line running across the chart represents the average tension of CO_2 during quiet normal breathing. The respiratory movements are represented by the tracing at the foot of the curve along the abscissa. It will be observed that the oxygen tension falls very rapidly during the apneic period, until just before breathing recommences it may be as low as 30-35 mm. Hg instead of the normal of about 95. Meanwhile the CO_2 tension rises from the very low level of 12 mm., at first very rapidly, then more gradually, although, when breathing recommences, it has not yet gained the normal level. As a result of the first periods of breathing, the O_2 tension suddenly shoots up, but the CO_2 falls only slightly. During the next apneic stage the O_2 quickly comes down again, and the CO_2 rises so as almost to attain normal tension before breathing again supervenes. As the apneic periods subsequently become less pronounced, the CO_2 tension comes to stand almost at its normal level, whereas considerable variations in the O_2 tension continue to occur.

Several interesting features of these results demand attention. In the first place, it is plain that the body is possessed of some mechanism by which it can prevent great fluctuations in the CO_2 tension of the blood, whereas towards O_2 no such "buffer action" is displayed. It will further be observed that the CO_2 tension of the alveolar air rises very rapidly during the first part of the apneic period, and then more gradually, the explanation being that during the forced breathing the CO_2 has been washed out from the blood but not from the body as a whole.

At first sight one might attribute the periodicity to the same cause as that operating during breathing through a long tube with soda lime—namely, to oxygen deficiency. But this explanation is untenable, because the periodicity remains evident for some time after all possibility of direct stimulation of the center of O_2 deficiency is over. A possible clue is furnished by the fact that breathing returns while the CO_2 tension is still considerably *below* its normal level. The return, as we have seen, is accounted for by the appearance of lactic acid, and if we assume that this has occurred particularly in the respiratory center itself, a slight degree of hyperpnea will be excited, which by supplying O_2 will

quickly oxidize the lactic acid, so that the still slightly subnormal C_H of the blood is unable to excite the center. Apnea therefore supervenes and lasts until lactic acid has again accumulated in the center. To explain why local accumulation of lactic acid in the center should produce a periodic type of breathing, we must further assume that there is considerable delay between the moment at which equilibrium of the gases in the blood and alveolar air becomes established and that at which the blood arrives at the respiratory center. This delay is caused by the slowing of the bloodflow on account of the absence of respiratory movements.

Emphasis is placed on the fact that it is in the center itself and not in the blood that the lactic acid becomes oxidized by the excess of O_2, because lactic acid is known to disappear slowly under these conditions from isolated blood, but to do so very quickly from tissues such as muscle, and presumably therefore also from nervous tissue.

In support of the above explanation it has been found that, if toward the end of the forced breathing the lungs are filled with sufficient O_2 so that the tension of this gas in the alveoli is not lower than 120 mm. Hg, breathing is regular in type when it returns, and the CO_2 tension of the alveolar air is several millimeters *above* instead of below the normal stimulating level.

To sum up, the periodic character of the breathing supervening on a period of apnea may be explained as follows: Under ordinary conditions of breathing and barometric pressure the O_2 tension of the blood is sufficient between normal respirations to prevent any accumulation of lactic acid in the respiratory center, so that the stimulus afforded by the C_H of the blood produces a constant effect. During the apnea which supervenes upon forced breathing, lactic acid accumulates in the center, causing this to respond to the gradually rising C_H of the blood before the latter has reached its physiologic level. The hyperpnea thus excited does not, however, bring about a prompt oxidation of the lactic acid in the center or a lowering of the C_H of the blood circulating through it, because more time than usual is taken for the blood to get from the lungs to the brain on account of the absence of respiratory movements. When the aerated blood does reach the respiratory center, the excess of O_2 which it contains oxidizes the lactic acid so that apnea supervenes, and the lactic acid again accumulates, although not now so much as before because of the gradually rising C_H of the blood itself. The essential factor in the causation of periodic breathing is therefore a delayed mass movement of the blood from the pulmonary capillaries to the respiratory center. The delay may be caused by cessation of the respira-

tory movement, as in postapneic periodicity, or by some pathologic circulatory disturbance.

Periodic breathing is produced by forced respiration more readily in rarefied air than at sea level. It was found by Douglas,[26] after breathing forcibly for one minute at sea level, that the breathing when it returned showed 8 to 10 different periods of apnea and hyperpnea. On repetition of the experiment at an altitude giving a barometric pressure of 600 mm., 25 such periods followed the apnea; at a height corresponding to 520 mm., 40 periods. Indeed, at high altitudes periodic breathing may be brought about by the slightest alteration in normal respiration; even taking a deep breath may be sufficient to cause distinct periodicity in the succeeding respirations, and in many persons living at high altitudes periodic breathing is very apt to occur during sleep. As in pathologic cases exhibiting Cheyne-Stokes respiration, the periodic breathing at high altitudes can be immediately removed by inspiring oxygen.

We have devoted considerable space to a discussion of these extremely difficult problems in the hope that clinical observers, by becoming acquainted with the purely experimental work, may be in a position to conduct more searching investigations as to the cause of Cheyne-Stokes and other pathologic forms of periodic breathing.

CHAPTER XLIV

RESPIRATION BEYOND THE LUNGS

Up to the present our studies in respiration have concerned the various mechanisms involved in bringing about a constant change in the composition of the alveolar air. We must now consider the nature of the means by which the oxygen is conveyed to the tissues and the CO_2 removed from them.

In the first place, it is important to note that it is not for purposes of oxidation in the blood itself that the O_2 is required. In its respiratory function this fluid serves as a transporting agency between the lungs and the tissues, in which reside the furnaces of the body that consume the O_2 and produce the CO_2. This does not imply that there is no oxidation in the blood itself; indeed, we should expect a certain degree of oxidation because of the fact that the blood contains some living cells—the leucocytes. It is scarcely necessary nowadays to offer evidence for the foregoing conclusion. One well-known experimental proof consists in *replacing the blood in a frog with physiologic saline solution* and then subjecting the frog with the saline in its blood vessels to an atmosphere of pure O_2, when it will be found that the animal continues to absorb the normal amount of O_2 and exhale the normal amount of CO_2. It respires normally without any blood in the blood vessels.

In order that this transportation of gases between the lungs and the tissues may be efficiently performed, the blood must be provided with means for carrying adequate amounts of gases to supply the requirements of the tissues, both during rest and during their varying degrees of activity. Not only, therefore, must the O_2 and CO_2 capacity of the blood be very considerable, but it must be capable of very rapid adjustment from time to time.

Our problem naturally resolves itself into three parts: (1) the call of the tissues for oxygen (Barcroft); or, as it is styled, tissue or internal respiration; (2) the mechanism by which the blood transports the proper amounts of gases to meet the requirements of the tissues; and (3) the mechanism by which the blood gases are exchanged in the lungs—external respiration. For convenience, however, we shall change this natural order and consider the transportation of the gases first.

THE TRANSPORTATION OF GASES BY THE BLOOD

The Transportation of Oxygen

It is plainly not by mere solution in the plasma of the blood that the transportation of O_2 occurs, for at the partial pressure of this gas existing in the alveolar air at the temperature of the body the amount that could be dissolved in the blood would be only one-fortieth of that which is actually found to be present. If there were only plasma in the blood vessels, it would require a volume of fluid amounting to 150 kilograms or more in order to convey the necessary amount of O_2 from the lungs to the tissues; that is, the contents of the vascular system would weigh twice as much as the average weight of a man.

The substance that carries the O_2 in the blood is the *hemoglobin,* which may be described as a highly complex iron compound of protein especially evolved for the purpose of transporting O_2. In some of the lower animals other compounds exist in the blood for this purpose, but none of them is to be compared in its efficiency with hemoglobin. They are merely poor imitations of it.

Regarding the conditions under which hemoglobin combines with or delivers up O_2, the first question that presents itself is whether or not the reaction is a strictly chemical one. If so, a definite amount of O_2 must be capable of combining with a definite amount of hemoglobin. It is impossible to secure hemoglobin of sufficient purity to test this relationship directly on hemoglobin itself, so that we must test it indirectly by examining the combining equivalent between O_2 and that portion of the hemoglobin molecule upon which the combining power depends. This is the part of the molecule containing iron. Now, if we compare the amount of O_2 which hemoglobin can take up with the amount of iron present in the hemoglobin, we shall find that *one atom of iron becomes combined with two atoms of O_2*. Evidently, then, we are here dealing with a definite chemical reaction occurring between the O_2 and the iron of the hematin portion of the hemoglobin. This relationship is known as "the specific oxygen capacity of hemoglobin."

In showing that the union of O_2 and hemoglobin occurs according to chemical laws, we throw into prominence consideration of the mechanism by which the O_2 combined with hemoglobin in the blood is rapidly delivered up in the capillaries so as to supply the tissues with their requirement, and is then as rapidly recombined again in the lungs. Moreover, we must reconcile facts implied by the idea of a specific O_2 capacity with the well-known observation that the hemoglobin in the circulation is usually united with considerably less O_2 than the total amount possible.

In other words, we must recognize that, although it is essentially a chemical reaction, the combination of O_2 with hemoglobin is greatly influenced by other factors, and that it is these that are likely to be of physiologic importance.

In order to understand *the conditions under which hemoglobin will take up and give off O_2 in the animal body,* we must study the combining power of hemoglobin when it is exposed to different partial pressures of O_2 (for laws governing this, see page 336). In the blood, the extremes of the partial pressure of O_2 are represented, at the one end, by that in the alveolar air, which we have seen to be about 100 mm. Hg, and at the other, by that existing in the tissues, such as muscle, which has been shown to be not more than 19 or 20 mm. Hg. We must further bear in mind that the O_2 in its passage from the alveolar air to the hemoglobin and from the hemoglobin to the tissues, is transmitted in solution through the plasma; that is, so far as the supply of O_2 to the tissue cells is concerned, the plasma serves as the immediate source. Since the tissues are using up O_2 at a very great speed, especially when active, and are thus tending to lower the tension of O_2 in the plasma, favorable conditions have to be created whereby the hemoglobin liberates O_2 at the same rate as that at which it is leaving the plasma. In brief, it is the *O_2 tension of the plasma in the tissue capillaries that is the important factor,* the hemoglobin merely serving as a storehouse, which delivers its O_2 at just such a rate as to maintain the plasma-oxygen tension at a constant level. It is obviously of the greatest importance that we should understand how this mechanism of an adequate plasma-oxygen tension is maintained.

Methods of Investigation.—We must remember that the combination of O_2 and hemoglobin, being a definite chemical reaction, will be reversible, and must, therefore, obey the laws of mass action (see page 23) according to the equation: $Hb + O_2 \rightleftarrows HbO_2$. In order to ascertain the position of the balance of this equation at different partial pressures of O_2,—that is, the relative quantities of oxy- and reduced hemoglobin formed in a solution of hemoglobin when this is shaken with O_2 at different pressures,—we may proceed as follows: A few c.c. of the hemoglobin solution are placed in each of a series of vessels called *tonometers,* like those shown in Fig. 134. In addition to the hemoglobin solution, each tonometer contains a mixture of nitrogen and O_2 in different proportions. Suppose we use six vessels and in No. 1 have pure nitrogen; in No. 2, nitrogen containing 5 mm. partial pressure of O_2; in No. 3, 10 mm.; in No. 4, 20; in No. 5, 50; and in No. 6, 100. We now rotate the tonometers in a water-bath at body temperature for about twenty minutes, so that, by the formation of a thin film of hemoglobin solution over

the walls of the vessel, perfect equilibrium between the atmosphere and the fluid may be attained (see page 338). A measured quantity of hemoglobin solution (0.1 or 1.0 c.c.) is then removed from each tonometer

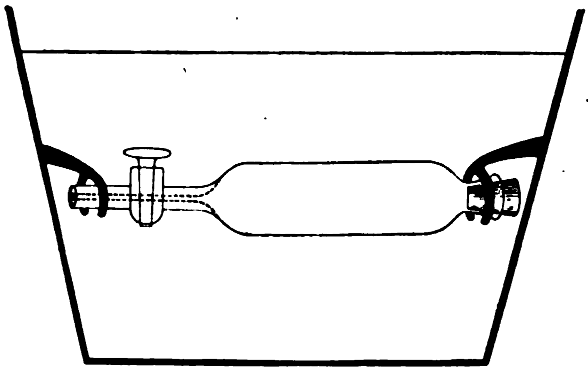


Fig. 134.—Barcroft's tonometer for determining the curve of absorption of oxygen by hemoglobin or blood. (From Starling's *Physiology*.)

and placed, together with some very dilute ammonia to lake the blood in one of the small bottles of the ***differential manometer,*** shown in Fig. 135.* This manometer consists in principle of a graduated U-shaped tube of narrow bore, containing clove oil, the free end of the U-tube

Fig. 135.—Barcroft's differential blood gas manometer. The capillary U-tube contains clove oil. The pockets on the sides of the blood bottles should be deeper. For manipulation see context.

being connected with small bottles provided with some device so that two fluids can be placed in each of them but kept unmixed until the bottle is violently shaken. The three-way stopcock between the small

*The blood-gas manometers are made in two sizes for use with 1 c.c. and 0.1 c.c. quantities of blood, respectively. The results with these small quantities are as accurate as with larger amounts.

bottles and the manometer serves to permit communication of the manometer with the outside air.

An equal quantity of hemoglobin solution that has been saturated with oxygen—i. e., oxyhemoglobin—is placed in the bottle on the other end of the manometer tube from that containing the bottle with the unsaturated hemoglobin solution. The bottles having been attached to the manometer with the stopcocks open to the outside, the apparatus is placed in a water-bath until the temperature conditions are constant. The manometers are then closed to the outside air and the bottles are shaken in order that the hemoglobin solution that is unsaturated with O_2 may take up O_2 from the atmosphere in the bottle until it becomes

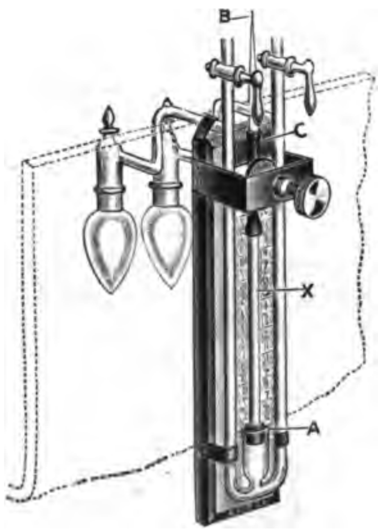


Fig. 136.—Barcroft blood gas manometer. This form can be used either as a differential manometer (page 390) or for direct measurement of pressure. For the latter purpose one bottle is removed and the pressure of gas generated in the other bottle is measured by the height to which it raises the clove oil in the distal tube of the manometer, the meniscus in the proximal limb being readjusted to its original level by compression with the brass screw of the rubber tube shown in the center.

saturated. The resulting shrinkage in the volume of the atmosphere on the side of the unknown hemoglobin solution causes the clove oil meniscus to move towards that side, the degree of movement being proportional to the initial unsaturation of the hemoglobin. The manometer tubes are then again brought into communication with the atmosphere so that the meniscus of clove oil may move back to its old level, and the bottle with saturated hemoglobin is removed from the manometer and a drop or two of a saturated solution of potassium ferricyanide placed in the separate compartment of the bottle without allowing it to mix with the hemoglobin. The bottle is then reattached, the temperature

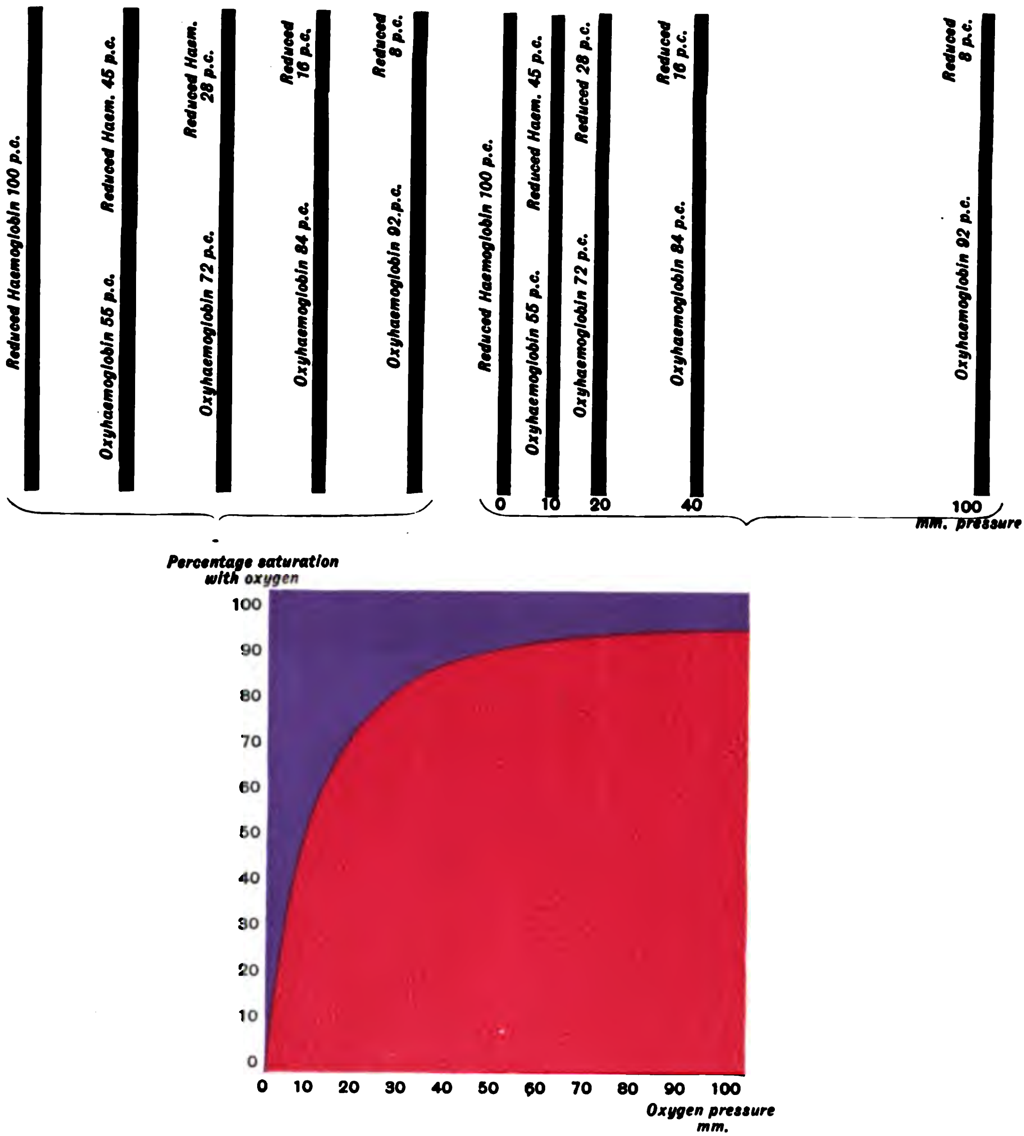


Fig. 137.—Upper left hand, percentage saturation of hemoglobin with oxygen at 37° C. corresponding to oxygen pressures of 0, 10, 20, 40 and 100 mm. of oxygen, respectively.
Upper right hand, the same spaced with the oxygen pressure as the abscissae.
Lower figure, dissociation curve representing the equilibrium between oxygen, oxyhemoglobin (red) and reduced hemoglobin (purple). (From Joseph Barcroft.)

conditions readjusted, the manometer closed off from the ouside air, and the apparatus again shaken so that the ferricyanide mixes with the hemoglobin solution. This drives off all the O_2 from the oxyhemoglobin solution, and, therefore, raises the pressure in the atmosphere of that bottle so that the clove oil moves to the opposite side of the manometer, the degree of displacement being proportional to the amount of oxyhemoglobin.

We have now all the necessary data for estimating the relative amounts of reduced hemoglobin in the hemoglobin solution as removed from the tonometers, for it is plain that the second estimation, as described above, tells us how much oxyhemoglobin might have been formed had all the hemoglobin been saturated and the first one, how much O_2 had yet to be taken up by the original hemoglobin solution to produce saturation.

The Dissociation Curve.—The next step is to plot the results obtained from the various hemoglobin solutions in the form of a curve. This is known as *the dissociation curve of hemoglobin.* It is plotted with the relative percentages of reduced and oxyhemoglobin in each of the solutions along the ordinates, and the partial pressures of O_2 in millimeters of mercury to which they were exposed along the abscissæ. The curve thus drawn is exactly of the same shape as that which would be produced if we were to place the tonometers in a row at distances from one another corresponding to the partial pressure of O_2 which each contained, and then to mark on each tonometer the relative amounts of reduced and oxyhemoglobin found in the solutions after shaking. A line joining these marks on the tonometers would then exactly correspond to the curve drawn by the method described above. This will be clear from the accompanying figure from Barcroft's book (Fig. 137).

In such a chart the space below the curve can be taken to represent the percentage of oxyhemoglobin (red in chart), and that above it of reduced hemoglobin (blue in chart), at the varying partial pressures of O_2 which are indicated along the abscissæ as being contained in the atmosphere of the tonometers, and which must be proportional to the partial pressure of O_2 in the solution in which the hemoglobin is dissolved.

Difference between Curves of Blood and Hemoglobin Solutions.—The curve obtained from *pure hemoglobin solutions* is very far, however, from clearing up the problem as to how the blood absorbs and discharges O_2. On the contrary, it makes this problem appear all the more difficult, for, according to the curve (Fig. 137) the hemoglobin is already more than half combined with O_2 at a partial pressure of this gas of no more than 10 mm. Hg, which means that in the low partial pressure of O_2 existing in the capillaries the oxyhemoglobin, in-

stead of readily yielding up its load of O_2, would greedily retain practically the whole of it. The curve, in other words, would satisfactorily explain why hemoglobin should readily absorb O_2 from the alveolar air, but would fall far short of explaining how this O_2 is readily released when it is required in the tissues. Obviously there is some artificial condition present in the above experiment which can not obtain in the natural environment of the blood.

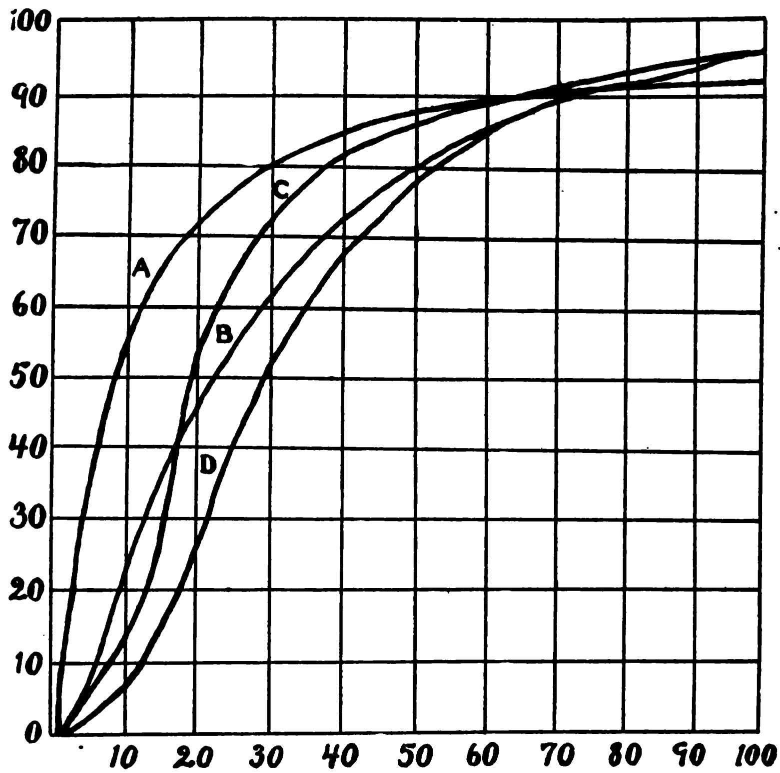


Fig. 138.—Average dissociation curves.

Ordinates—Percentage saturation of hemoglobin with oxygen.
Abscissæ—Tension of oxygen in mm. of mercury.
Curve A—Degree of saturation of pure hemoglobin solutions at varying pressures.
Curve B—Disregard this curve.
Curve C—Effect of 20 mm. CO_2 pressure on above solution.
Curve D—The saturation curve in normal blood at 40 mm. carbon dioxide pressure.

Since hemoglobin takes up O_2 in proportion to its iron, it can not be because of changes in the O_2 combining part of the hemoglobin itself that blood and pure hemoglobin solutions have dissimilar dissociation curves, but rather because of differences in the environment in which the hemoglobin acts. That this is so can be readily shown by plotting the dissociation curve, not for a hemoglobin solution, but for *blood* itself

(*D* in Fig. 138). The results are very different. At a partial pressure of O_2 of about 60 mm. Hg—that is, a lower pressure than exists in the lung alveoli (100 mm.)—the blood becomes nearly saturated with O_2, whereas at pressures below 50 mm. it readily loses O_2, so that at 10 mm. there is nearly complete reduction.

The question is: What are the environmental conditions under which the hemoglobin in the blood so alters its combining power for O_2 as to

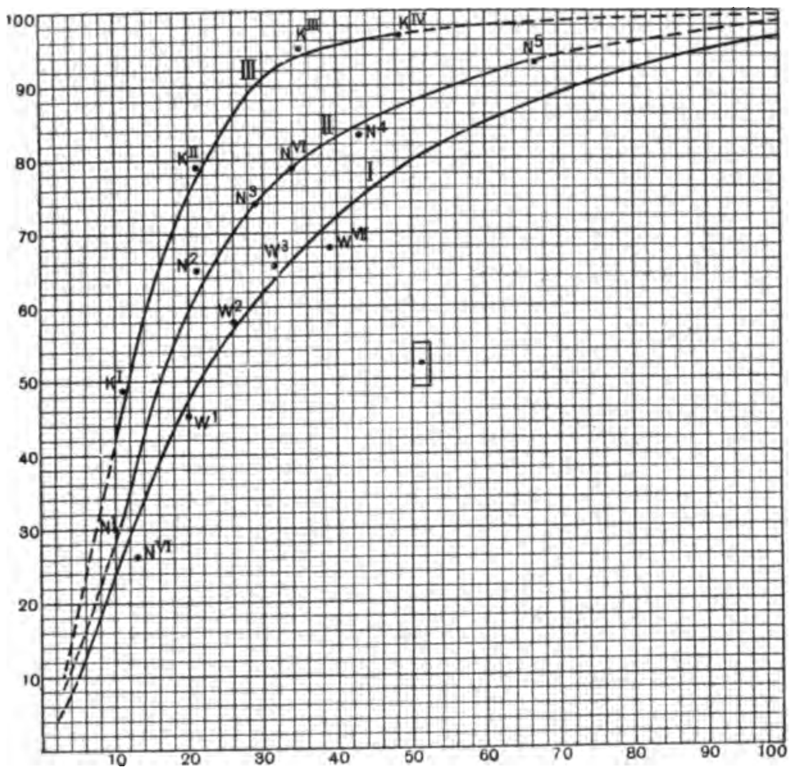


Fig. 139.—Dissociation curves of hemoglobin.

Ordinates—Percentage saturation of hemoglobin.
Abscissa—Tension of oxygen in mm. of mercury.
I. Dissociation curve of hemoglobin dissolved in water.
II. Dissociation curve of hemoglobin dissolved in 7% NaCl.
III. Dissociation curve of hemoglobin dissolved in 9% KCl.
Temperature 37-38° C. (From Joseph Barcroft.)

produce such a difference in the dissociation curve? By experimenting with hemoglobin solutions, three such factors have been found to come into play: (1) the presence of inorganic salts, (2) the hydrogen-ion concentration (CO_2 tension) of the solution, and (3) the temperature. If hemoglobin is dissolved in water containing the various *salts* of plasma in the same proportion as in blood (artificial plasma), the dissociation curve will be found to change so as to resemble that of blood (Fig. 139).

Since the plasmas of different animals contain different proportions of salts, the artificial plasma required to secure the result is not always the same. It differs, for example, for the dog and man. Potassium salts are particularly efficient in causing hemoglobin to absorb O_2. The influence of varying *hydrogen-ion concentrations* of the solution may be conveniently studied by adding varying percentages of CO_2 to the gas mixture in the tonometers, when it will be found that the curve becomes lowered in proportion to the amount of CO_2 present. This is shown in Fig. 140.

The effect of *temperature* on the dissociation curve is twofold: (1) on the rate with which equilibrium is established at the given partial pres-

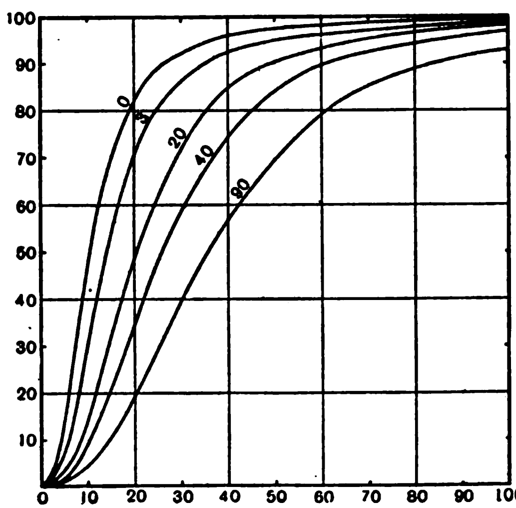


Fig. 140—Dissociation curves of human blood, exposed to 0, 3, 20, 40 and 90 mm. CO_2. Ordinate, percentage saturation. Abscissa, oxygen pressure. (From Joseph Barcroft.)

sure of O_2, and (2) on the position of the curve; the lower the temperature, the higher the curve.

The Rate of Dissociation.—Though it is now clear that the three conditions—namely, saline content, C_H, and temperature—are capable of altering the dissociation curve of a pure hemoglobin solution so as to make it correspond with that of blood, this does not entirely solve our problem, for we have yet to show how the cooperation of these forces renders it possible for the rate at which hemoglobin takes up O_2 in the lungs to correspond exactly with that at which it gives up its O_2 to the tissues. To study this problem a somewhat different kind of experiment must be undertaken. The hemoglobin solution is placed in a tube and the gas mixture slowly bubbled through it, samples of the solution being removed at intervals for analysis in the differential blood-

gas apparatus. To obtain the rate of oxidation, a mixture of N_2 or H_2 and O_2 is bubbled through the blood with the partial pressure of the O_2 the same as that which obtains in alveolar air—namely, about 95-100 mm. Hg; and to obtain the rate of reduction pure N_2 or H_2 gas is bubbled through.

The *rates* of reduction or of oxidation as thus determined are then plotted in curves constructed with the percentage saturation of the

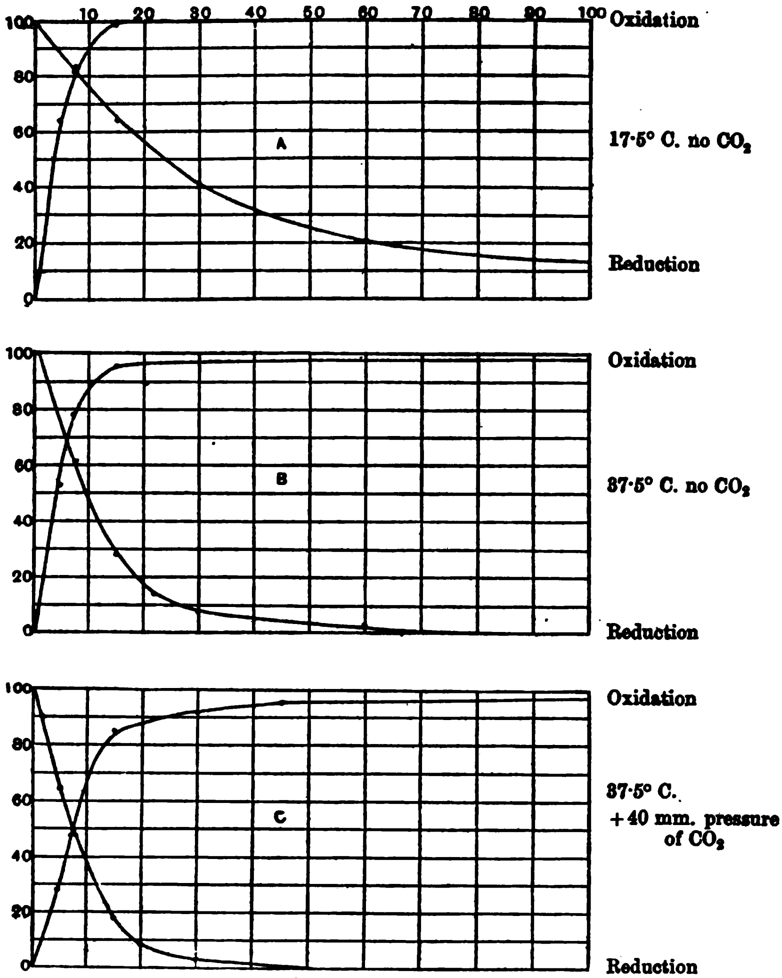


Fig. 141.—Curves showing relative rates of oxidation and reduction of blood as influenced by temperature and tension of CO_2.
Ordinates—Percentage saturation.
Abscissae.—Time in minutes.
Reducing gas, hydrogen.
Oxidizing gas, oxygen.
A, temperature 17.5° C., with no CO_2.
B, temperature 37.5° C., with no CO_2.
C, temperature 37.5° C., but the O_2 and H contained 40 mm. Hg pressure of CO_2. (From Joseph Barcroft.)

hemoglobin on the ordinates and the time in minutes along the abscissæ (Fig. 141). Even if we use blood in this experiment and therefore make certain that the hemoglobin is acting in the presence of the proper proportion of salts, we shall find, as Fig. A shows, that at room temperature the rate of oxidation is very much greater than the rate of reduction. If now we repeat the observation at a temperature of 37° C., the two curves come more nearly to correspond, but still the rate of reduction is slower than that of oxidation. If in a third experiment, besides having proper temperature and chemical conditions, we produce the oxidation and reduction in the presence of a partial pressure of CO_2 of 40 mm., which corresponds to that of the arterial blood, we shall find that oxidation becomes a little slower, whereas reduction is further quickened. Indeed the two curves, as seen in *C* in the figure, come practically to correspond, indicating that the environmental conditions under which hemoglobin combines and gives off O_2 in the blood are exactly adjusted.

One word more with regard to the influence of C_H. Its effect in flattening out the curve, especially at the lower partial pressures of O_2, indicates that when a high C_H is present, the blood will very readily part with its O_2 supply. Now, the most significant application of this fact is that high concentrations of H ion will occur just exactly where it will be of benefit—namely, in the capillaries (because of the CO_2 and lactic acid produced by the tissues). Some doubt has, however, recently been thrown on the importance of this factor.

Since, as we have seen, hemoglobin absorbs O_2 according to chemical laws, it will naturally be asked not only why the dissociation curve flattens out while yet maintaining the shape of a right-angled hyperbola, as by the action of acids or an increase in temperature, but also why it should change its shape when salts are also present. The explanation offered by Barcroft and his pupils is that the changes depend on the fact that hemoglobin being a colloidal substance, its molecules undergo processes of aggregation under the conditions referred to above, and therefore cause the reaction to become of a different type from that represented by the equation $HbO_2 \rightleftarrows Hb + O_2$. As has been pointed out by Bayliss, although such an explanation might suffice to explain the flattening out of the curve, it fails to explain the change in its shape; for, according to the laws of mass action, such a change could occur only if molecules of a different type came to take part in the reaction.

Dissociation Constant.—Notwithstanding these criticisms, it is of considerable practical importance to know that an equation exists from which the entire dissociation curve can be plotted by making only *one* determination of the relative amounts of oxy- and reduced hemoglobin at a particular tension or partial pressure of oxygen. This equation is as

follows: $\frac{y}{100} = \frac{Kx^n}{1+Kx^n}$, where y equals the percentage saturation of hemoglobin with O_2, x the O_2 tension, and K and n are constants, K being the equilibrium constant and n the average number of molecules of hemoglobin supposed to exist in each aggregate.

When this equation is applied to human blood, the value of n remains unchanged and is given as 2.5, so that by transposition we are enabled to find the value of K as follows: $K = \frac{y}{x^n(100-y)}$. If we find the value of K by measuring the relative saturation of the blood with O_2 at *one* pressure of this gas, then by changing the value of x to correspond to other O_2 pressures, we can find all positions of the curve for a given sample of blood.

An important practical application of this method is found in *the determination of the C_H of blood,* for, as we have seen, the dissociation curve becomes lowered in proportion to the concentration of hydrogen ions. The acidity of a sample of blood can therefore be found by comparison of its dissociation curve, as plotted from the values found for K, with that of normal blood to which known quantities of acid have been added. When the curves correspond, the bloods must contain the same amounts of acid, other things being equal. In brief, then, *the reaction of the blood is proportional to the value of K.* When this is low, it indicates that the blood is taking up an abnormally low percentage of its possible load of O_2 at a given pressure of O_2, and that the acidity is greater than normal; when K is high, for the same reason the acidity must be low.

In determining K for the blood as it exists in the body, it is necessary that it should be subjected to the same tension of CO_2 as obtains in the blood vessels. K will then be proportional to the C_H of the living blood. This condition would be impossible to fulfil in drawn samples were it not for the fact that we can place in the tonometer an atmosphere containing the same partial pressure of CO_2 as is found in the alveolar air. Since this value varies in different individuals, it must be separately ascertained in each case (see page 344). As determined with these modifications, K has been found to vary in healthy men between 0.000212 and 0.000363 (ten individuals). When acid substances appear in the blood, as in acidosis, K becomes extremely low; thus, in one case suffering from acidosis with dyspnea, it was found a few hours before death to be only from 0.000082 to 0.00011. Similarly K becomes low in the acidosis associated with mountain sickness, and it is said to be raised after taking food that is rich in alkali.*

*When K is found to be normal, the blood is said to be *mesectic;* where K is low, it is said to be *myonectic;* and when K is high and the acidity is therefore small, it is said to be *pleonectic.*

CHAPTER XLV

RESPIRATION BEYOND THE LUNGS—Cont'd

THE MEANS BY WHICH THE BLOOD CARRIES THE GASES

In the foregoing account of the physiology of the blood gases, emphasis is placed on the tension under which the gases exist rather than on the total amount of each gas present in the blood. This has been done because the exchange of gases between alveolar air and blood and between blood and tissues proceeds according to the laws of gas diffusion, which are of course dependent upon differences in gas pressure or tension.

Something must now be said regarding the *amount of the gases.* This may be measured either by physical or by chemical methods. In the former, a measured quantity of blood is received into an evacuated glass vessel, which is then attached to a mercury pump, by which the gases are sucked out of the blood and transferred, by suitable manipulations of stopcocks, to a graduated tube, in which they are then analyzed by chemical means. The principle of the chemical method has already been described in connection with the measurement of oxygen in hemoglobin solutions (see page 382). A measured quantity of blood, kept free from contact with the air, is transferred under some weak ammonia solution to one of the blood-gas bottles of the blood-gas differential manometer, and a few drops of a saturated solution of potassium ferricyanide is placed in the pocket of the bottle. After the blood has been laked and temperature conditions adjusted, the ferricyanide is mixed with the blood solution, thus causing the O_2 to be quantitatively displaced. From the increased pressure produced in the manometer the amount of O_2 can readily be computed. To determine the CO_2 of the blood, the bottle is now removed from the manometer and a few drops of a saturated solution of tartaric acid placed in the pocket. When this is mixed with the deoxygenated blood mixture, after the usual adjustment for temperature, the pressure caused by the evolved CO_2 is recorded and the amount present calculated.

The results of the analysis are expressed as the number of cubic centimeters of gas present in 100 c.c. of blood—the volume percentage, as it is called. The following are approximate percentage values:

	OXYGEN	CARBON DIOXIDE	TOTAL GAS
Venous blood	12	48	60
Arterial blood	20	40	60

The estimation of the amounts of the gases, although of little value in connection with the physiology of gas exchange, is very important in supplying information regarding the respiratory activities of the various organs and tissues. Just as we determine the total respiratory exchange of an animal by measuring the differences in O_2 and CO_2 in inspired and expired air, so may we determine the degree of tissue respiratory exchange by analysis of the gases in blood removed from the artery and vein of the tissue. It should be clearly understood, however, that it is not the percentage but the total amount of the gases that must be considered, and that it is therefore necessary to know the volumes of blood-flow as well as the percentage of the gases. Something will be said later of the results of such investigations (see page 393).

At present we are concerned with the manner in which gases are carried in the blood. The O_2, as we have seen, is carried by the hemoglobin, some being also in a state of simple solution in the plasma. The CO_2, which it will be noted is present even in arterial blood in considerably greater amount than the O_2, is partly combined with alkali to form bicarbonates. The alkali available for this purpose varies from time to time according to the amount of other acid substances present. Since these are stronger acids than carbonic, any increase in their amount (acidosis) causes displacement of some of the CO_2, thus bringing about, as we have seen, a relative increase in free CO_2 in the blood and therefore raising the C_H.

What particularly interests us here is *the agency by which the combined CO_2 is carried in the blood.* If blood is exposed to a full atmosphere of CO_2, it will take up as much as 150 per cent of the gas—that is, between two and three times the amount ordinarily present in it. It has therefore a great reserve capacity for CO_2. A greater proportion of the CO_2 is carried in the plasma than in the corpuscles; but if plasma (or serum) is exposed in a vacuum, all of the CO_2 present in it will not be evolved. When blood itself is similarly exposed, on the other hand, all the CO_2 is given off. To liberate all of the CO_2 from plasma *in vacuo,* some acid must be added, from which it has been inferred that blood corpuscles act like weak acids. It is commonly stated that hemoglobin or some constituent of blood is capable of freeing CO_2 from solutions of sodium carbonate, but the recent work of Buckmaster[32] shows that this is not the case. The decomposing power of blood is caused by the development of acidity in the shed blood and any similar power that the corpuscles may exhibit is due to a discharge from

them into the plasma of acid radicles. At least it has been found that the alkalinity of the plasma increases when CO_2 is bubbled through blood, this increase in alkalinity being interpreted as the result of the migration of acid radicles into the corpuscles. This would lead us to expect that under the opposite conditions (i. e., *in vacuo*) acids would leave the corpuscles.

Proteins are amphoteric substances—that is, they combine with acids or alkalies—which would lead us to expect that they would be capable of absorbing some CO_2. That this is the case, particularly for hemo-

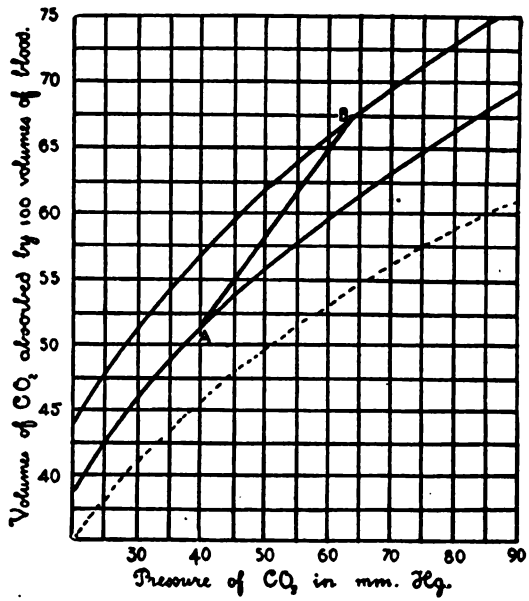


Fig. 142.—Curve of CO_2 tension in blood. For description, see text. (From Christiansen, Douglas and Haldane.)

globin, has been shown by comparing the CO_2-combining powers of water and a solution of pure hemoglobin.

Attempts have been made to determine the relative amounts of CO_2 carried by these various agencies in the blood. The following is an example of such a table:

In simple solution in plasma and corpuscles			1.9 c.c.
As sodium bicarbonate	α in corpuscles	6.8	18.8 "
	β in plasma	12.0	
In combination with hemoglobin		7.5	19.3 "
In combination with proteins of plasma		11.8	
			40.0

(Loewy.)

The power of blood to absorb CO_2 at various tensions of this gas, as determined in the same way as for O_2 (see page 380), has shown that saturation of the hemoglobin with O_2 distinctly diminishes the CO_2-carrying power of the blood. This is shown in the accompanying curves

The various tensions of CO_2 are given along the abscissæ and the volume per cents of CO_2 taken up by the blood on the ordinates. The upper curve is drawn from results obtained when the blood was shaken with CO_2 in the presence of hydrogen, and the lower, when in the presence of air. (The dotted curve may be disregarded.) The line *AB* drawn between the two curves represents the absorption of CO_2 by the blood within the body. At a tension of 40 mm. CO_2—that present in alveolar air (see page 356)—*A* stands in arterial blood at about 52 vols. per cent; and at a pressure of 62 mm.—possibly present in the tissues—*B* stands in venous blood at about 67 vols. per cent. The CO_2-containing power would be 7 per cent lower (i. e., 60 vols. per cent) in blood saturated with O_2 at the latter pressure. The oxygenation of blood in the lungs, therefore, helps to drive out the CO_2; and conversely, its deoxygenation in the tissues enhances its power of absorbing this gas.

Having shown how the blood transports its charge of O_2 from the lungs to the tissues, we may now proceed to study the call for O_2 by the tissues, and in this connection we have to consider (1) the amount of O_2 which they require under varying conditions of rest and activity, and (2) the mechanisms by which their varying demands are met.

THE OXYGEN REQUIREMENT OF THE TISSUES

In order to ascertain the average O_2 requirement of the different tissues of the body, it is necessary to adopt as a standard of measurement the amount of O_2 in c.c. absorbed per gram of tissue per minute. To obtain it we must know: (1) the weight of the particular organ or tissue under investigation; (2) the bloodflow through the vessels of the organ in c.c. per minute; and (3) the different percentages of O_2 in the arterial and venous blood of the tissue. It would be beyond the scope of this book to review in any detail the many experimental investigations which have been undertaken in this connection. A few of the most recent and important results are given in the accompanying table from Halliburton's *Physiology*:

In the order of their oxygen requirements, or *the coefficient of oxidation,* as it is called, the tissues may be divided into four groups; glandular, muscular, connective, and nervous. The nervous tissues should possibly stand above the connective, but very little is known regarding their oxygen consumption, although it appears that this is quite low (Hill and

ORGAN	CONDITION OF REST	OXYGEN USED PER MINUTE PER GRAM OF ORGAN	CONDITION OF ACTIVITY	OXYGEN USED PER MINUTE PER GRAM OF ORGAN
Voluntary muscle	Nerves cut. Tone absent	0.003 c.c.	Tone existing in rest Gentle contraction Active contraction	0.006 c.c. 0.020 c.c. 0.080 c.c.
Unstriped muscle	Resting	0.004 c.c.	Contracting	0.007 c.c.
Heart	Very slow and feeble contractions	0.007 c.c.	Normal contractions Very active	0.05 c.c. 0.08 c.c.
Submaxillary gland	Nerves cut	0.03 c.c.	Chorda stimulations	0.10 c.c.
Pancreas	Not secreting	0.03 c.c.	Secretion after injection of secretin	0.10 c.c.
Kidney	Scanty secretion	0.03 c.c.	After injection of diuretic	0.10 c.c.
Intestines	Not absorbing	0.02 c.c.	Absorbing peptone	0.03 c.c.
Liver	In fasting animal	0.01 to 0.02 c.c.	In fed animals	0.03 to 0.05 c.c.
Suprarenal gland	Normal	0.045 c.c.	——	——

Nabarro). It is of course necessary in making these comparisons to secure the coefficient of oxidation both when the tissue is at rest and when it is thrown into varying degrees of activity. Special attention has been devoted to the requirements of skeletal muscle, heart muscle and the salivary glands.

Skeletal Muscle.—In observations on skeletal muscle, Verzár (cf. 27) isolated the gastrocnemius muscle of the cat, and without disturbing its blood supply collected samples of blood by introducing a 1 c.c. pipette into a branch of the saphenous vein. Activity was produced by throwing the muscle into tetanus by the application of an electrical stimulus to the sciatic nerve. During its contraction the muscle lifted a weight, so that it did about 70 gram-centimeters of work at the beginning of each period of tetanus. The velocity of bloodflow was determined by the rate at which the blood flowed along the pipette, and the O_2 consumption, by the difference in percentage of O_2 in the venous and the arterial blood. These measurements were made: (1) before contraction, (2) during contraction, and (3) after contraction. It was found that during the tetanus the O_2 consumption in some cases was greater than during rest, while in others it was actually less, but in every instance a great increase in O_2 consumption followed the tetanus—that is, the call for O_2 continues for some time after the actual work has been performed. This result

shows that the contraction is not dependent upon oxidation, but that *the oxidation occurs after the contraction is over.* The mechanism involved in muscular contraction can not therefore be analogous with that by which energy is liberated in a steam engine by the oxidation of the coal. The mechanism must rather be like that of a spring, which becomes unwound during the muscular contraction and requires O_2 for its rewinding.

Interesting results corroborative of these conclusions have been secured by observations on the *heat production* of isolated muscles. It was found that heat production occurred after a single shock to the muscles, not only during the contraction, but for a considerable period after it, provided O_2 was present. In the absence of O_2 this recovery was either greatly delayed or entirely abolished. Such results favor the view that O_2 is used largely in the processes whereby the muscles, "like an engine charging an accumulator, synthesize substances containing a considerable amount of potential energy, which again, like the accumulator, it discharges when appropriate stimuli are applied"—(L. V. Hill, cf. 27). One immediately thinks of lactic acid in connection with these interesting results, for, as has already been stated, Hopkins and Fletcher[29] have shown that this acid is produced in the absence of O_2 in excised frog muscles, but when O_2 is present, it is either not produced or, if so, quickly disappears.

Heart Muscle.—Another muscle that has been thoroughly investigated in this connection is that of the heart. The gaseous exchange has been studied both on isolated heart preparations and by examining the exchange in the lungs of a combined lung and heart preparation. The most important investigations by the first of these methods are those of Rohde (cf. 27), who arrived at the very important conclusion that the O_2 taken in by the heart muscle varies directly with the maximal tension set up in the heart by the contraction. This tension was measured by placing a rubber bag in the ventricle and distending it with water at a known pressure. By altering the initial pressure and by observing the pulse rate, it was found that the O_2 used by the heart depends on the product of the pulse frequency and the maximal increase in pressure produced by each cardiac contraction; or, in the form of an equation: $\frac{Q}{NT}$ = a constant quantity; where Q is the oxygen used, T the maximal increase of pressure at each beat, and N the frequency of the pulse.

It should be pointed out, however, that constancy in the product of the above equation does not hold under abnormal conditions of the heartbeat. For example, when the pressure in the heart is very high, the amount of O_2 required begins to go up out of proportion, indicating that

the heart is becoming overtaxed—that it is losing its efficiency. The same result occurs when the heart is dying, and when depressing drugs are used, such as chloral hydrate, potassium cyanide, veratrine, etc. Some other drugs, however, such as epinephrine, do not cause alteration in the ratio, nor does vagus stimulation. Of course when the vagus is stimulated, the O_2 consumption in a given period decreases because the heartbeats are slowed; but the absorption of O_2 is not increased relatively to the slowing of the heart.

Glands.—Most work has naturally been done on the most accessible gland—the submaxillary. By stimulating the secretory nerve of this gland (the chorda tympani) in the dog, it has been found that, whereas the more abundant secretion lasts only so long as the stimulus is applied to the nerve, the O_2 consumption is increased to several times that of rest, and remains increased for a considerable period after the stimulus has been removed. Accompanying the increased functional activity in such structures as muscles, there is a very marked increase in bloodflow due to vasodilatation, which, in part at least, is dependent upon the secretion into the blood of some substances resulting from the glandular activities, and is not entirely due to the action of vasodilator nerve fibers.

Similar results have been obtained in the case of the *pancreas* when excited to secrete by the injection of secretin (see page 425). Under such conditions, the oxygen consumption has been observed to increase about fourfold and to be accompanied by a dilatation of the gland.

The work on the *kidney* has been especially interesting, because it has been found that increased activity, which of course is measured by the rate of urine excretion, is not always accompanied by increased consumption of oxygen. When diuresis is produced by injecting Ringer's solution into the circulation, a great increase in urine outflow may occur without any change in oxygen consumption; whereas, on the other hand, when a diuretic such as sodium sulphate or caffeine is used, the oxygen consumption increases enormously.

Regarding the other tissues and organs, the O_2 consumption of the lungs and brain appears to be small. It is a very significant fact, however, that the higher cerebral centers are extremely sensitive to deprivation of O_2.

The Blood.—In the blood itself, a certain amount of oxidation goes on because of the presence of leucocytes. This oxidation becomes considerable in the blood of animals rendered anemic by the injection of phenyl hydrazin. A thorough investigation of the cause of this greater oxidation has shown it to be owing, not to an increase in nucleated corpuscles, but to the presence of the young unnucleated red blood

corpuscles, which appear in large numbers in the blood under these conditions. A similar increase in blood oxidation occurs during posthemorrhagic anemia, the rate of oxidation running parallel with the rate of regeneration of the red corpuscles.

The Mechanism by Which the Demands of the Tissues for Oxygen Are Met

There are two possible methods by which this may be brought about: (1) by a change in the C_H or the saline constituents or the temperature of the plasma, so that the hemoglobin more readily delivers up its load of O_2; and (2) by an increase in the mass movement of blood through the vessels of the acting tissue.

Regarding the first of these possibilities, there is no doubt that acids are produced during metabolism of acting tissues. As we have seen, when muscles contract in the presence of an abundance of O_2, CO_2 is produced in large amounts, and when they contract in a deficiency of O_2, sarcolactic acid. In the submaxillary gland, too, it has been possible to show that the C_H of the venous blood, as measured by the value of K of the dissociation curve of hemoglobin, becomes distinctly increased during glandular activity. That this increase in C_H will dislodge O_2 we have already seen (page 386). As to the possible influence of local changes in temperature and in saline constituents of the plasma, nothing can at present be said.

Regarding the second possibility, vasodilatation may be dependent either upon the action on the blood vessels of nerve impulses coming along vasomotor nerves, or upon the production by the active tissue of vasodilating or depressor substances (see page 243). Much evidence has been accumulating in recent years which tends to show that such depressor substances are produced, and they may be either (1) acids, or (2) organic bases of a similar nature to β-imidazolylethylamine (histamine). This latter substance is of considerable physiologic interest because of its close relationship to one of the main amino acids of the protein molecule—namely, histidine (see page 604). Its effect in producing vasodilatation is extraordinary. Thus, half a milligram of the drug injected intravenously into a monkey will lower the mean arterial pressure by fifty per cent.

But before such an hypothesis can be entertained, it is necessary to show that, independently of nerve impulses, the blood vessels of an acting organ may dilate. The best evidence has been secured by studying the effects of stimulating with epinephrine the cervical sympathetic nerve to the submaxillary gland of a cat. The gland cells become more active,

and dilatation of the artery occurs, although on blood vessels alone epinephrine in similar dosage produces constriction. Of course in showing that local chemical products of activity serve as the excitant of local dilatation, we do not mean to imply that the vasodilator fibers going to the blood vessels are of no use. Indeed we know that such fibers do become active in the case of a salivary gland whose cells have been paralyzed by atropine, but it is a significant fact that this dilatation is of relatively short duration, whereas that produced by glandular activity lasts for some time. The suggestion seems therefore not out of place that under normal conditions the initial dilatation of an acting gland may be brought about through nervous stimuli, but the later dilatation is maintained by metabolic products.

CHAPTER XLVI

THE PHYSIOLOGY OF BREATHING IN COMPRESSED AIR AND IN RAREFIED AIR

In the application of a knowledge of the physiology of respiration to the investigation of disease, a group of conditions arises in which considerable interference with physiologic mechanisms occurs, not as a result of disease, but of changes in the atmospheric environment. The regulation of the functions of respiration depends very largely on changes in the physical and chemical properties of the alveolar air, so that it is to be expected that similar changes in the atmosphere will have a marked influence on the respiratory activity and on the general well-being of the animal.

The most thoroughly investigated of these conditions are those which develop in rarefied and compressed air. Either condition can be produced experimentally in the laboratory by the use of air-tight chambers (pneumatic cabinets) and suitable pumps, although most of the important work on the effects of rarefied air has been conducted at high altitudes, where the barometric pressure is low.

MOUNTAIN SICKNESS

This condition depends primarily on disturbances in the control of the respiratory function, and it is on account of the useful information concerning the nature of these functions, rather than because of the so-called disease itself, that so much attention has been devoted to its investigation during recent years. The disturbances produced by the rarefied atmosphere develop rather quickly, but after some time they gradually disappear, indicating that the organism has acclimated itself—that is, the compensatory mechanisms have come into play to bring the respiratory control back to normal. When animals are placed in pneumatic cabinets from which some of the air is pumped out, most of the immediate symptoms observed in mountain sickness occur, but it is usually impracticable to continue the observations for a sufficient length of time to allow the compensating mechanisms to develop.

Because of their great value in revealing the nature of the respiratory hormone, many of the results of the recent investigations on mountain

sickness have been given elsewhere in this volume (page 360), where the general symptoms are also described. In this place we shall consider very briefly some of the more *general aspects of the condition,* and, more particularly, the nature of the adaptation that occurs. All of the symptoms are essentialy dependent upon lack of oxygen. Cyanosis is common and the symptoms are much the same as those of coal-gas poisoning. Not only does this deficiency of oxygen cause acid substances to appear in the blood, thus raising the C_H and stimulating the respiratory center, but it allows other poisonous materials to accumulate. These act on the various nerve centers, producing symptoms which vary in different individuals according to their relative susceptibilities. In some, the digestive centers are affected and nausea and vomiting occur; in others, the higher cerebral centers are affected, causing depression and general mental apathy, great drowsiness, muscular weakness, or it may be mental excitement and loss of self-control.

The susceptibility of different individuals also varies according to the amount of previous experience in mountaineering and the type of breathing. Much of the value of previous experience and training depends on the ability to perform muscular effort economically; to adjust the effort to the available oxygen supply without permitting unoxidized harmful products to accumulate in the body. It often happens that no symptoms appear so long as the person is at rest, but immediately do so whenever any muscular effort demands a much more abundant oxygen supply.

The type of breathing that best withstands the rarefied air is slow and deep, rather than rapid and shallow. The reason for this is of course that much more of the outside oxygen gets into the alveoli in the former case than in the latter, the dead space being practically constant. The following figures taken from observations on three different individuals will illustrate the importance of this factor.

	C.C. PER RESPIRATION	NO. OF RESPIRATIONS PER MINUTE	HEIGHT IN METERS AT WHICH SYMPTOMS OCCURRED
Subject 1	270	20	3300
" 2	440	14	6000
" 3	700	8	6500

(From Halliburton.)

After living for some time in the rarefied air and quite independently of training in the efficient performance of muscular work, *adaptation* occurs, so that the symptoms pass off. The essential feature of this adaptation is increased absorption of O_2 into the blood. Three mechanisms have been described as responsible for this effect: (1) increase in the tension of O_2 in the alveolar air; (2) assumption by the pulmonary epithelium

of the power of secreting O_2 into the blood; (3) increase in the erythrocytes and hemoglobin of the blood. The increased alveolar O_2 tension is a result of the more rapid breathing brought about by the increased C_H of the blood. If no adaptation occurred, the O_2 tension at 10,000 feet would be 59 mm. and at 15,000 feet, 33.8 mm. Actual observations on men, however, gave at 10,000 feet a tension of 65 mm. and at 15,000 feet, 52 mm.

The evidence for an increased secretory activity of the pulmonary epithelium depends on observations made by Haldane and his coworkers,[33] who found that blood collected from the finger of a man living on a high mountain is brightly arterial, whereas if this same blood is shaken in a flask with alveolar air from the man from whom it was taken, it will become darkly venous. To account for this difference it is believed that the pulmonary epithelium forces O_2 into the blood contrary to the laws of diffusion.

A more exact proof was sought for by comparing the relative amounts of O_2 and CO that blood would take up (1) when exposed outside the body and (2) while in the blood vessels. Carbon monoxide has a very great avidity for hemoglobin, so that if blood is shaken in a flask with air containing 0.07 per cent of this gas, colorimetric measurement will show an equal mixture of oxy- and carboxy-hemoglobin. Since carbon monoxide is destroyed with extreme slowness in the body, it is possible by causing a man to breathe a mixture of it in air to determine, in a sample of drawn blood, whether as much carboxy-hemoglobin has been formed as *in vitro*. If so, the O_2 tension in the blood must equal that in the alveoli; if less carboxy-hemoglobin should be formed, it would indicate that a higher tension of O_2 exists in the blood. This latter is the result which Haldane states he has secured. In one experiment, for example, when blood was shaken outside the body with 0.04 per cent CO_2, the amount of carboxy-hemoglobin formed was 31 per cent of the whole hemoglobin. When the same mixture was inhaled for three or four hours the percentage of carboxy-hemoglobin in the blood rose only to 26 per cent, which would correspond to an O_2 tension of 25 per cent of an atmosphere, whereas even at sea level the tension of O_2 in the alveolar air can not be above 15 per cent of an atmosphere.

The constant low tension of O_2 in the plasma stimulates the red blood corpuscles and the percentage of hemoglobin to become markedly increased after residence for some time in high altitudes. At first this is due to a concentration of the blood by a diminution in plasma, but gradually the blood-forming organs become excited and an actual increase in the total amount of hemoglobin occurs. In the light of these facts it is interesting to compare the average number of red corpuscles in the blood of inhabitants living at different altitudes.

	HEIGHT ABOVE SEA (METERS)	RED CORPUSCLES (PER C.MM. BLOOD)
Christiania	0	4,970,000
Zurich	412	5,752,000
Davos	1560	6,551,000
Arosa	1800	7,000,000
Cordilleras	4392	8,000,000

(From Starling.)

COMPRESSED-AIR SICKNESS; CAISSON DISEASE; DIVER'S PALSY

Divers and caisson workers are susceptible to peculiar symptoms. These are frequently of sufficient severity to cause death, but may be so mild as almost to escape notice. They first appear, not when the worker is subjected to the high pressure, but after he has come back to atmospheric pressure.*

While in the compressed air the worker as a rule suffers no discomfort. A stuffiness may be felt in the ears and temporary giddiness; the respiration and pulse rate may become slow and frequency of micturition may be noticed, but none of the symptoms of disease appear until after the caissonier or diver has been decompressed (after he has returned to atmospheric pressure), the exact time of their onset being either immediately after decompression or at the end of several hours. The worker may have returned home and spent the evening feeling perfectly well until he went to bed, when symptoms supervened which may include muscular and joint pains, vertigo, embarrassed breathing, subcutaneous emphysema and hemorrhages, pains in the ears and deafness, vomiting, perhaps hemoptysis and epigastric pain. These symptoms usually pass off after some hours but the arthralgia and myalgia sometimes persist for a considerable time.

In the more severe cases the first symptom is severe pain in the muscles and joints, quickly followed by motor paralysis, so that the patient falls and is likely to become unconscious. The pulse is almost imperceptible, the respiration is labored, sometimes even asphyxial, the face cyanosed, and the surface of the body cold. Many of the cases are fatal; indeed, death may be almost instantaneous. Such cases are common in careless diving when the divers, to return the more quickly, screw up the outlet valve in their helmets so as to fill their suits with air, which car-

*A caisson is a steel or wooden chamber sunk in water and prevented from filling by means of compressed air. For the passage of the workmen and of material, into and out of the caisson, the latter is connected with a second smaller chamber fitted with air-locks and decompressing cocks. A diver works in a waterproof suit, the head being enclosed in a copper helmet connected by hose with air pumps. Every 10 meters or 33 feet of water corresponds to one atmosphere pressure (15 pounds to the square inch), so that at this depth the total air pressure in a caisson, or in a diver's helmet, would amount to 30 pounds to the square inch, that is, + 1 atmosphere.

ries them to the surface, where they decompress themselves by opening the valve.

Autopsies of persons dead of caisson disease have shown, as a rule, intense congestion of the viscera, hemorrhages in the spinal cord and brain, and ecchymoses on the pleura and pericardium. In some cases interlobar emphysema of the lungs and laceration of the spinal cord and brain have been noted.

The Cause of the Symptoms

The cause for the symptoms is not, as was at one time supposed, that the pressure drives the blood from the peripheral into the deep regions of the body, including the nerve centers. Such a process is impossible, because the fluids of the body—and all tissues, even the bones, are full of fluid—are incompressible. Pressure applied to any part of the body will be immediately distributed equally to every other part. If this were not so, life would be impossible during *any* variation of atmospheric pressure. It is now clearly established that all the symptoms of caisson disease are due to *decompression*, and not, in the slightest degree, to the mechanical effect of the pressure itself (Paul Bert, Leonard Hill and Macleod[34]).

When an animal is under pressure, its tissue fluids dissolve a large amount of gas. They absorb it in obedience to the law of solution of a gas in a fluid, which states that the amount of gas dissolved in water is directly proportional to the partial pressure of that gas in the atmosphere; at two atmospheric pressures twice as much gas will pass into solution as at zero pressure (Dalton's law). So long as the gas is in simple solution, it does not in any way change the physical condition of the blood and tissue fluids. If, however, the animal is suddenly decompressed (i. e., the pressure of air surrounding it is reduced to zero), the dissolved gas will be so quickly thrown out of solution that bubbles of it are set free. These bubbles act as air emboli, sticking in the pulmonic capillaries or blocking up a terminal artery in the brain; or they may be large and tear the capillary wall and so lead to hemorrhage. If these bubbles are produced in the posterior spinal roots, intense pain results; if in the anterior, motor paralysis. Frothing of the blood in the heart impedes the action of the organ and death soon follows.

The following experiments furnish proof of this explanation: A frog was placed in a small steel chamber connected with a cylinder of compressed air and provided with two windows by which a strong arc light could be passed through the chamber. The web of the foot was stretched on a wire and fixed so that the small blood-vessels could be seen by applying a microscope to the outside of the window. After carefully observing the circulation of the blood in the vessels at atmospheric pressure, a posi-

tive pressure, amounting in some experiments to + 50 atmospheres, was introduced but no effect could be noted on the circulating blood. By opening a tap in the chamber, decompression to zero pressure was quickly effected and, immediately, large bubbles were seen to develop in the blood, blocking the vessels and producing stasis. The bubbles were derived from the gas that had gone into solution under pressure. On reapplying the pressure the bubbles of gas again went into solution and the blood circulated normally. When the pressure was subsequently very gradually lowered to zero, the circulation went on undisturbed, and the frog was removed from the chamber in normal condition.

The process involved in causing caisson disease is evidently the same as that which can be observed in a bottle of aerated water; if the cork in such a bottle is drawn, the dissolved gas escapes as bubbles and effervescence results; if the bottle is recorked, the gas reenters solution and the fluid becomes quiet. If a pin hole is made in the cork, the gas will gradually escape and no effervescence will result.

Confirmatory results have been secured by observations on mammals. The arterial blood pressure of rabbits was not found to become altered by exposure to compressed air, and various animals placed in a large, strong steel chamber at pressures far in excess of those to which man ever subjects himself did not show any symptoms like those of caisson sickness, unless the pressure was suddenly lowered. Many times also, if symptoms had appeared they could be removed by again subjecting the animals to the compressed air.

Investigations were also carried out to determine exactly how much gas the blood of an animal subjected to high pressures contains, and how long it takes to absorb the maximal amount of gas and to release it. It was found that the gases that increased in amount were nitrogen and oxygen, and that these become dissolved in the blood according to Dalton's law.

The Prevention of the Symptoms

The most important practical application of these observations concerns the length of time required for the saturation and desaturation to occur, for *the results serve as a basis upon which the safe regulation of work in compressed air by man can be conducted.* The most significant outcome of the above experiments from this standpoint is that it takes considerable time for the blood to absorb its full quota of gas at a given atmospheric pressure and to liberate it again when the animal is decompressed. The cause of delay is that the tissue fluids other than the blood take much longer than would be expected to reach equilibrium with the partial pressure of gas in the blood plasma.

To understand why this delay should occur, let us suppose that the only gas concerned is nitrogen. As the pressure rises, the blood in the capillaries of the lungs must dissolve nitrogen in proportion to the pressure of this gas in the alveoli; the blood carries the dissolved gas to the tissues and these dissolve it until the pressure is again equalized between them and the blood. The blood, after giving up its excess of dissolved nitrogen, returns to the lungs and again becomes saturated and this goes on until blood and tissue have become saturated with gas at the external pressure. The tissues are two-thirds water and they contain (in man) from 15 to 20 per cent of fat. Fat, however, dissolves five times more nitrogen than water (Vernon); consequently, it takes longer for a given volume of tissue than of blood to become saturated at a given pressure.

The blood in man constitutes one-twentieth of the body weight; so that if the tissues were all liquid they would dissolve 20 times as much nitrogen as the blood. On account of the fat which they contain, however, the tissues take up more than this proportion—namely, in an average man about 35 times more than the blood. All the blood in the body takes about one minute to complete a round of the circulation, so that in this time, after being suddenly subjected to an increased pressure—assuming that the blood circulates equally throughout the body—the tissues will be one-thirty-fifth saturated; in the next minute another thirty-fifth of thirty-four thirty-fifths will be saturated, and so on. After five minutes the body will be about 22 per cent, and in 25 minutes about one-half, saturated; but *it will take about two hours before saturation is complete.* These calculations assume that the blood is evenly distributed throughout the body; but this is not the case, for its mass movement varies considerably in different parts, being much greater in the active muscles and in the glands than in passive structures, such as fat. These less vascular parts will therefore lag behind the others in taking up their full quota of gas, and therefore prolong the time necessary for complete saturation of the body as a whole.

We see therefore that, after some time in compressed air, the blood and active tissues will be saturated and contain volumes of dissolved gas in proportion to their relative bulks; the fat, although not saturated, will yet contain up to five times more gas than an equal volume of blood, and the passive tissues will be incompletely saturated.

These considerations regarding the saturation of the different parts of the body apply also in its desaturation. Suppose, for example, that the external pressure is suddenly lowered: the blood, on leaving the lungs, will contain no excess of gas; when it reaches the tissues it will remove gas until the pressure is equalized, discharge this into the alveoli and return again for more. Other things being equal, it will take the

same number of minutes to desaturate that it took to saturate, and the parts of the body that will lag behind the others, in being desaturated, are those with a sluggish circulation.

When the mass movement of the blood is increased by muscular exercise, the rate of saturation and desaturation with nitrogen is increased in proportion. During active work the increase in movement of the blood may be four or five times over the normal, so that the tissues of the caisson worker become much more quickly desaturated during decompression than the above figures would lead one to expect.

Application of Foregoing Laws in Practice

With regard to the application of these principles in the decompression of caisson workers, it is impracticable to occupy as much time as it takes to saturate the body even at comparatively low pressures. If the great dangers attending work in compressed air are to be avoided, we must either insist on very gradual decompression or we must show how the dissolved gases may be got rid of by some modification in the decompression procedure. With this object in view, we must determine what difference of pressure may be allowed between the external air and the body without the formation of bubbles. Actual experience shows that there is no risk of bubble-formation, however quick the decompression, after exposure to + 15 pounds pressure (i. e., 2 atmospheres absolute). "Now, the volume of gas capable of being liberated on decompression to any given pressure is the same, if the relative diminution of pressure is the same"—(Haldane[35]). On reduction from 4 to 2 atmospheres, the same volume of gas will tend to be liberated as on reduction from 2 to 1 atmospheres—that is to say, no bubbles will form. The practical conclusion is "that the absolute air pressure can always be reduced to half the absolute pressure at which the tissues are saturated without risk." Thus, after saturation at 90 pounds absolute pressure (+ 5 atmospheres), a man can be immediately decompressed to 45 pounds (+ 2 atmospheres) in a few minutes without risk, but from this point on the decompression must be conducted slowly, so as to insure that the nitrogen pressure in the tissues is never more than twice the air pressure. The great advantage of this method is that it makes the greatest possible use of difference of pressure between tissues and blood in order to get rid of the gas that these contain.

When the decompression from the start is gradual, the desaturation of the tissues will progressively lag behind that of the blood, and the tendency to the liberation of free gas will become greater. In such a case the decompression is far too slow at first and far too rapid later.

Theoretically, therefore, *the decompression should be rapid at first and very slow later.*

Before recommending the adoption of this principle of stage decompression in caisson work, Haldane and his coworkers made numerous observations on the incidence of decompression symptoms in laboratory animals. They assert that the stage method is decidedly safer than the uniform method, the advantage being particularly after short exposures. On the other hand, Leonard Hill could make out no definite advantage for the stage method. The two methods have also been compared in actual caisson work at the Elbe Tunnel, where the pressure was + 2 atmospheres. Very little advantage could be demonstrated for the stage as compared with the uniform method at this comparatively low pressure. The general conclusion which we may draw is that the stage method should be employed, although it is not to be expected that it will absolutely insure absence of decompression symptoms. Of course the great advantage of the stage method is the saving of time, making it possible to persuade the workmen to adopt it.

There are two other factors that are to be considered in hastening the desaturation of the tissues; these are *muscular exercise*, and the *breathing of an indifferent* gas.

It is clear, from what has already been said, that the gas dissolved in the tissues will become removed in proportion to the mass movement of the blood, and it is probably true that muscular exercise, performed in the decompression chamber, is of as great importance in preventing the subsequent development of symptoms as a much prolonged decompression. In a man at rest, the circulation through the central nervous system and the viscera is constantly influenced by the pumping action of the respiratory movements, but in the capillaries of the muscles, joints, fat, etc., this influence is not felt and the blood flows more slowly. It is consequently in these parts that bubble formation is likely to occur, especially some time after decompression. The bubbles cause the neuralgic pains—the "bends" and "screws" so well known to caisson workers. These could no doubt be entirely prevented by muscular exercise and massage of the limbs during decompression. In illustration of these facts the following experiment by Greenwood may be cited: During decompression from + 75 pounds pressure in 95 minutes "Greenwood flexed and extended all the limb joints at frequent intervals, with the exception of the knees. Subsequently pain and stiffness were experienced in the knees and nowhere else." In another experiment the knees also were flexed and no pain was felt.

But even in the parts with active circulation, the gas in the tissues may lag considerably behind that in the blood, although the decompres-

sion has been properly controlled. This has been shown by Leonard Hill in the case of the kidney. The "tissue" gas in this case can be taken as the gas dissolved in the urine, by analyzing which, therefore, at different stages of decompression, the excess of nitrogen over what it should be at the external pressure, can be ascertained. On decompression from + 30 pounds by two stages to zero, a considerable supersaturation was found to exist. The excess of nitrogen can, however, be cleared out of the kidneys rapidly and completely by breathing oxygen, which should therefore be administered during decompression in cases where great care has to be exercised (Leonard Hill).

When symptoms do appear, they can, in most cases, be relieved by recompression, and all modern caisson works are provided with a special chamber for this purpose. We need scarcely say anything about this treatment here, as its value is so well known. Suffice it to say that, although it is most likely to afford relief when applied as soon as possible after the appearance of the symptoms, yet it is often efficacious when applied several days after their onset.

Quite apart from the dangers of decompression, it must of course be remembered that the working conditions in a caisson are somewhat different from those at atmospheric pressure, as the air, owing to its compression, is warmer and is loaded to saturation point with moisture. This hot, wet air interferes with the heat-regulating mechanism of the body, making hard muscular work very uncomfortable because of the tendency of the body temperature to rise. The reaction of the body against this tendency to hyperthermia consists in dilatation of the superficial capillaries and increased heart action.

When such working conditions are repeated day by day, the appetite is likely to fail, partly because of the tendency of the body to suppress the activity of the metabolic processes, so as to keep down heat production, and partly, no doubt, because the digestive processes are working below par on account of there being less blood circulating through the visceral blood vessels, it having been sent to the surface of the body to be cooled off. The worker therefore tends to take less food, his metabolism becomes depressed, and his factors of safety against bacterial infections become lessened.

The risk of the appearance of symptoms on decompression is also greater when the air in the caisson has been moist and hot, for the heart has been overworking to maintain the bloodflow in the dilated vessels; it gets fatigued and is consequently unable to maintain, during decompression, a rate of bloodflow that is adequate for carrying the gas-saturated blood to the lungs, where the excess of gas becomes dissipated.

The criterion of proper working conditions in the caisson is therefore the wet-bulb temperature. This should stand below 75° F. To maintain this condition it is necessary to ventilate the caisson, preferably with air that has been cooled by cold-water radiators; in any case, the ventilation should be adequate to keep down the wet-bulb temperature. The increased expense of ventilation with cooled air would soon be balanced by the greater working efficiency of the men. Constant circulation of the air in the caissons by means of fans assists also in improving the conditions, for it helps to increase dissipation of heat from the body.

CHAPTER XLVII

THE CIRCULATORY AND RESPIRATORY CHANGES ACCOMPANYING MUSCULAR EXERCISE*

During activity the muscles require many times more blood than during rest. When the activity is widespread the greater blood supply is provided by increased heart action accompanied by dilatation of the muscular arterioles and constriction of those of the splanchnic area, so that the entire available blood supply of the body is made to circulate more rapidly. When, on the other hand, the activity is confined to a limited group of muscles, the increased blood supply is mainly provided by a local dilatation of the blood vessels of the active muscles accompanied by a reciprocal constriction of those of inactive parts. Under these conditions there may therefore be no quickening of the bloodflow as a whole. In order that this accurate adjustment of blood supply to tissue demands may be promptly and adequately brought about, all available types of coordinating mechanism are called into play; that is to say, mechanical, nervous and hormone factors cooperate to an extent which is dependent upon the type of work being performed.

Besides the changes in pulse rate and blood pressure which are evidently designed to supply more blood to the acting muscles, changes dependent upon a secondary effect of the muscular movements have also to be considered. Although the various factors work together and are more or less interdependent, the final effect can be understood only after we have studied the relative influence of each separately.

The Mechanical Factor.—It is particularly with regard to this factor that the circulatory changes may be an unavoidable consequence of, rather than a useful adjustment to, the muscular effort. The effects vary with the type of exercise performed. In repeatedly lifting and lowering dumbbells from the floor to above the head, the contracting muscles of the back and extremities and of the abdomen compress the veins and cause the blood to flow more rapidly into the heart, so that the arterial pressure suddenly rises. So long as this compression exists, the veins remain relatively empty and the arteries overfilled, but whenever it ceases and the muscles relax, the veins fill up again and the arterial pres-

*This chapter is placed here rather than following circulation because of the interdependence of the circulatory and respiratory adjustments.

sure markedly falls, until the extra space in the veins has been occupied by blood. It is for this reason that the arterial blood pressure is always found to be little, if any, above normal when taken within a few seconds after such exercise. It subsequently rises because the other factors responsible for the increased pressure (quick heart and arteriole constriction) are still in operation at the time the veins again become filled with blood. The purely mechanical influence outlasts the exercise for a comparatively short time, whereas the nervous and hormone influences continue acting. This interpretation is supported by the observation that the fall of blood pressure is greater when the subject is left standing after a given amount of dumbbell exercise than when he is allowed to sit with his elbows resting on his knees. In the standing position the pressure on the abdominal veins is less and the hydrostatic effect of gravity causes more blood to collect in the large veins (Cotton, Rapport and Lewis[36]). Being purely mechanical in its causation, the preliminary fall following dumbbell exercise can always be demonstrated if the observations are made at close enough intervals of time.

The mechanical response of the circulation to exercise acts therefore through the rate of filling of the right heart with blood, and if this organ is in a healthy condition, it will respond to the greater inflow by correspondingly increased discharge. Like every other physiologic mechanism, the heart works with a large factor of safety—a reserve power—and it is the rate of venous filling that determines how much of this reserve must be called upon to maintain the circulation. In isolated heart-lung preparations Starling and his coworkers have very clearly demonstrated the close dependence of cardiac output upon rate of venous filling and the enormous range through which the systolic discharge can be made to vary by altering this factor. As explained elsewhere, when the reserve power of the heart is lessened, the rise in blood pressure following exercise is longer in attaining its maximum, which is set at a higher level and persists for a longer time. Observation of the extent of these changes furnishes a most useful functional test of cardiac efficiency.

Other mechanical factors that augment the cardiac output depend on the increased respiratory movements. During each respiration the increase in capacity in the thorax causes both an opening up of the thin-walled veins, so that blood is aspirated towards them from the extra-thoracic venous system, and a dilatation of the blood vessels of the lungs, so that the blood finds its way from right to left heart more readily. Although this dilatation will at first tend to cause more blood to collect in the intrathoracic vessels and less to be pumped out of them, the expiratory act when it supervenes will, by compressing the veins, cause the extra blood to be expelled into the left ventricle and thence into the

arteries. It is obvious that increased depth and frequency of the respiratory movements will accelerate the bloodflow and tend to raise the arterial blood pressure.

The above factors will come into play during most kinds of muscular exercise such as walking, running, or swinging dumbbells, etc. There are certain types of muscular effort, however, in which the mechanical factors produce decidedly *disturbing effects* on the circulation. During a sustained effort as, for example, in pulling against a resistance or in attempting to lift a heavy load, the respirations are suspended, often after a deep inspiration, and the contracted abdominal muscles press the diaphragm up into the thoracic cavity. After a preliminary squeezing out of blood first of all from the veins of the abdomen into the thorax and then from those of the latter into the systemic arteries, with a consequent rise in arterial pressure, there comes to be a damming back of blood into the peripheral veins, causing them to swell and, if continued, marked cyanosis may develop. When such efforts are maintained for long, the arterial pressure begins to fall, and this fall is very pronounced indeed at the end of the effort, because, the compression being removed from the abdominal and thoracic veins, these open up and form a large unfilled blood reservoir.

A similar mechanism comes into play during expulsive acts such as defecation, parturition, etc. In these the glottis is closed, usually after a preliminary inspiration, and a powerful expiratory movement is performed, with the consequence that the intrathoracic and intraabdominal pressures rise considerably, greatly augmenting the systolic discharge and causing the blood pressure to rise. Because of the obstruction to the bloodflow in the large veins of the abdomen and thorax, however, the later effect of the effort is to diminish the systolic discharge, but the fall in blood pressure which this would be expected to occasion is masked. The pressure remains high because other factors increasing the peripheral resistance come into play. The fall in blood pressure following these acts may be very marked indeed. Similar mechanical effects are produced in the acts of coughing, sneezing, etc.

The capacity of the veins varies considerably with the position of the body, and it is in order that we may cause alterations in this capacity and therefore encourage a more rapid bloodflow that we stretch the body after sitting for some time in a cramped position.

The Nervous Factor.—The vagus, vasoconstrictor and respiratory centers are all excited during muscular effort. In the earlier stages the excitation depends entirely on nervous impulses transmitted to the centers, but later it depends on changes in the composition and temperature of the blood flowing through them—the hormone factor. The initial

stimulation of the centers must be due to cerebral impulses independently transmitted to the above centers, since the quickening of the pulse and respirations may be observed to begin before the actual muscular contractions.

The Hormone Factor.—We have to consider first the nature of the hormone, and secondly the mode of its action.

The Nature of the Hormones.—The most important hormone is carbonic acid, but when the exercise is strenuous and continued, or from the very start is of such a nature that it uses up oxygen more quickly than the blood can supply it to the muscles, lactic acid also appears. Evidence for these statements can readily be supplied in man by analysis of the expired air (for carbon dioxide) and of the urine (for lactic acid) before and during muscular work. The real hormone in both cases is believed to be an increase in the H-ion concentration of the blood. There is, however, no direct proof of this assertion—that is to say, no one has actually shown that a measurable change in the H-ion concentration of the arterial blood (for of course a change in the venous blood would be of no significance) does occur before the changes believed to be dependent upon acid production make their appearance. The well-known buffer action of the blood (that is, its ability to take up considerable quantities of acid or of alkali before any perceptible change occurs in H-ion concentration) furnishes another reason why doubt must be cast upon the H-ion hypothesis. The most delicate means for demonstrating a change in H-ion concentration of the blood consists in finding the dissociation constant for hemoglobin and the results have shown that acidosis develops during exercise at least at high altitudes (Barcroft[1]). So far as we are aware, however, it has not been possible by direct measurement (page 29) to detect a rise in H-ion concentration. Of course it may well be that the sensitiveness of the various nerve centers and other structures towards the H-ion concentration is very much greater than our most refined and sensitive laboratory methods can reveal. Such is at least commonly believed to be the case for the respiratory center (see page 351), and it may also be so for those of vascular tone and cardiac action. It is nevertheless possible that an increase in the free carbonic acid itself—the carbonate anion ($-HCO_3$), in other words—is the effective hormone. In the first stages of muscular work, this increase would be due to greater production of CO_2, whereas later, especially when the work is strenuous, lactic acid would decompose the $NaHCO_3$ of the blood, liberating $-HCO_3$, which would become added to that still being produced by the active muscles, and as the $NaHCO_3$ (buffer substance) became gradually used up, would cause a relatively greater and greater proportion of $-HCO_3$ to exist in a free

state. That the CO_2 tension of the alveolar air should be found to be lowered by prolonged muscular exercise in no way detracts from this explanation, for it is dependent upon the greatly increased rate of movement of air into and out of the alveoli (see also page 366).

One serious difficulty in accepting the HCO_3 ion as the exciting hormone of the nerve centers during muscular exercise depends on the observation that the alveolar CO_2 after some time is lower than normal. If we accept Haldane's teaching that there is accurate correspondence between the tensions of CO_2 in arterial blood and alveolar air not only during rest but also during muscular activity, then obviously we must discard the HCO_3 hypothesis. Leonard Hill and Flack,[37] however, have shown quite clearly both in experimental animals and in man that equilibrium between the blood and alveolar tensions of CO_2 may fail to occur. When blood with excess of CO_2 is injected into the jugular vein of dogs, the respiratory center is stimulated, as shown by the increased breathing, which indicates that the CO_2-rich blood must have passed through the lungs without the excess of CO_2 being removed from it. Hill believes that the diffusion of CO_2 out of the blood into the alveolar air may be depressed in muscular exercise, and that this rather than the appearance of lactic acid in the blood is responsible for the low CO_2 tensions usually found present (see page 369). He points out in support of this view that a person after exercise can hold his breath for a much shorter time than is usual, and the CO_2 meanwhile mounts in the alveolar air very rapidly.

The only way by which progress may be made in a problem like that under discussion is, however, to adopt some hypothesis and then to gather evidence for or against it. At the present stage of our knowledge, the hypothesis usually adopted is that a slight change in H-ion concentration of the blood is the effectual hormone. It is an hypothesis which is supported by the parallelism between the effects observed during muscular exercise and those produced by experimental increase in H-ion concentration.

The Effects of the Hormone.—These may be classified as follows: (1) strictly local effects on the muscles themselves; (2) effects on the heart; and (3) effects on the nerve centers. The local production of acids in the muscles will cause dilatation of the arterioles, for it has been shown by various observers that acids cause relaxation of vascular muscle. Even the capillaries themselves are said to be dilated by carbonic acid (Severini). The effects produced on the heart by changes in H-ion concentration of the blood have been particularly studied by Starling and Patterson,[38] who, working on isolated heart-lung preparations, have shown that the heart relaxes more and more and discharges less blood

as the H-ion concentration of the perfusion fluid is increased by adding CO_2 to the air ventilating the lungs.

The influence of changes in H-ion concentration of the blood on the vagus and vasomotor centers is usually believed to be stimulatory. There is no doubt that an increase in C_H stimulates the vasoconstrictor centers, not only of the medulla, but also, although much more feebly, of the spinal cord. But it is a question whether any part of the rise in systolic pressure during muscular exercise can be attributed to this cause, for the enormously increased bloodflow which is known to occur makes it problematical whether any vasoconstriction really occurs. If it does so, it must be confined to the splanchnic area, where it would have the effect of bringing about a redistribution of the total available blood by expressing it from the viscera and sending it to the active muscles.

The effect of increased H-ion concentration on the vagus center must be insignificant. It is commonly believed that it would cause not what is actually observed, a quickening, but rather a slowing of the heart rate. But even this is doubtful. The slowing of the heart that is observed in asphyxia, for example, is in part at least due to the increased intracranial pressure, for when the carotid artery is connected with a mercury valve so that the blood escapes as the pressure rises above the normal level, no slowing of the heart is said to occur in asphyxia. As Leonard Hill and Flack[37] have shown, however, a part of the slowing is due to the direct effect of CO_2. If increase in the H-ion concentration does affect the heart during muscular exercise, it must act by inhibiting the vagus tone, which is opposite to the action which it is usually believed to have. The activity of the respiratory center is of course excited by increase in H-ion concentration, and this, as we have seen, will cause important changes in the circulation because of the mechanical effects which follow.

Along with hormones we must consider the effect of change in the *temperature of the blood.* That this rises during muscular exercise is well known, but that it should be responsible for many of the cardiovascular adjustments that occur is quite commonly overlooked. It is, for example, very likely that rise in blood temperature is responsible for the acceleration of the heart that occurs during exercise when both vagi have been severed, and it no doubt is responsible for a part at least of the vasodilatation and respiratory acceleration.

Finally, it is interesting to speculate as to the nature of the changes that occur when the "second wind" is acquired during strenuous muscular exercise. In running, for example, considerably more distress is experienced a short time after the start than some time later. Three

very definite changes occur at the time the relief is experienced—namely, a slowing and steadying of the previously much quickened and irregular pulse, sweating, and a marked fall in the respiratory quotient. The last mentioned change possibly gives a clue to the cause of the others. In the early stages *R. Q.* is raised, which indicates that relatively more CO_2 is being expelled from the blood into the alveolar air than oxygen is being absorbed, perhaps because of inadequate movement of blood through the lungs. At the time of the adjustment it is possible that a pronounced vasodilatation occurs in the muscles and coronary arteries. The former change by lowering the arterial blood pressure will relieve the pumping action of the heart, and the latter will improve its power of contraction by supplying it with more oxygen.

RESPIRATION REFERENCES

(Monographs)

Barcroft, J.: The Respiratory Function of the Blood, University Press, Cambridge, 1914.
Borrutau, H.: Nagel's Handbuch der Physiologie, 1905, i, 29.
Douglas, C. G.: Die Regulation der Atmung beim Menschen, Ergebnisse der Physiologie, 1914, p. 338.
Hill, Leonard: Caisson Sickness, International Medical Monographs, E. Arnold, London, 1912.
Keith, Arthur: The Mechanism of Respiration in Man, Further Advances in Physiology, E. Arnold, London, 1909.
Schenck, F.: Innervation der Atmung, Ergebnisse der Physiologie, 1908, p. 65.

(Original Articles)

[1]Keith, Arthur: Cf. Further Advances.
[2]Hoover, C. F.: Arch. Int. Med., 1913, xii, 214; ibid., 1917, xx, 701.
[3]Lee. F. S., Guenther, A. E., and Meleney, H. F.: Am. Jour. Physiol., 1916, xl, 446.
[4]Meltzer, S. J.: Jour. Physiol., 1892, xiii, 218.
[5]Haldane, J. S., and Priestley, J. G.: Jour. Physiol., 1905, xxxii, 225.
Haldane and Douglas: Ibid., 1913, xlv, 235.
[6]Henderson, Y., Chillingworth and Whitney: Am. Jour. Physiol., 1915, xxxviii, 1.
Henderson and Morriss: Jour. Biol. Chem., 1917, xxx, 217.
[7]Krogh, A., and Lindhard: Jour. Physiol., 1913, xlvii, 30; ibid., 1917, li, 59.
[8]Pearce, R. G.: Am. Jour. Physiol., 1917, xliii, 73; ibid., 1917, xliv, 369.
[9]Siebeck, R.: Skand. Arch. f. Physiol., 1911, xxv, 87; Carter, E. P.: Jour. Exper. Med., 1914, xx, 21.
[10]Peabody, F. W., and Wentworth, J. A.: Arch. Int. Med., 1917, xx, 443.
[11]Lewis, T.: Jour. Physiol., 1908, xxxiv, 213, 233.
[12]Porter, W. T.: Jour. Physiol., 1895, xvii, 455.
[13]Christiansen and Haldane, J.: Jour. Physiol., 1914, xlviii, 272.
[14]Boothby, W. M., and Berry, F. B.: Am. Jour. Physiol., 1915, xxxvii, 433; also Boothby, W. M., and Shamoff, V. N.: Ibid., p. 418.
[15]Alcock, N. H., and Seemann, J.: Jour. Physiol., 1905, xxxii, 30.
[16]Scott, F. H.: Jour. Physiol., 1908, xxxvii, 301.
[17]Stewart, G. N., and Pike, F. H.: Jour. Physiol., 1907, xx, 61.
[17a]Coombs, H. C., and Pike, F. H.: Proc. Soc. Exper. Biol. Med., 1918, xv, 55.
[18]Krogh, A.: Skand. Arch. f. Physiol., 1910, xxiii, 248; and A. Krogh with Marie Krogh, ibid., 179.

[19]Haldane, J. S., and Priestley, J. G.: Jour. Physiol., 1905, xxxii, 225.
[20]Scott, R. W.: Am. Jour. Physiol., 1917, xliv, 196.
[21]Newburg, Means, and Porter, W. T. : Jour. Exper. Med., 1916, xxiv, 583.
[22]Hasselbalch, K. A., and Lundsgaard, Chr.: Biochem. Ztschr., 1912, xxxviii, 77, and Skand. Arch. f. Physiol., 1912, xxvii, 13.
[23]Hooker, D. R., Wilson, D. W., and Connett, H.: Am. Jour. Physiol., 1917, xliii, 357.
[24]Campbell, J. M. H., Douglas, C. G., and Hobson, F. G.: Jour. Physiol., 1914, xlviii, 303.
[25]Lindhard, J.: Jour. Physiol., 1911, xxxviii, 337; Haldane, J. S., and Douglas, C. G.: Ibid., 1913, xlvi.
[26]Douglas, C. G.: Art, Ergebnisse der Physiologie, *see* Monographs.
[27]Barcroft, J.: *see* Respiratory Function of Blood.
[28]Milroy, T. H.: Quart. Jour. Physiol., 1913, vi, 373.
[29]Fletcher, W. M., and Hopkins, F. G.: Jour. Physiol., 1907, xxxv, 247; also Fletcher, W. M.: Jour. Physiol., 1913, xlvii, 361.
[30]Ryffel, J. H.: Proc. Physiol. Soc. in Jour. Physiol., 1909, xxxix, 29.
[31]Pembrey, M. S., and Allen, R. W.: Jour. Physiol., 1909, xxxii, 18.
[32]Buckmaster, G. A.: Jour. Physiol., 1917, li, 105.
[33]Douglas, C. G., Haldane, J. S., Henderson, Y., and Schneider, E. C.: Phil. Trans. Roy. Soc., 1913, 203, B, 185.
[34]Hill, Leonard, Macleod, J. J. R.: Jour. Physiol., 1903, xxix, 507; Hill, Leonard, Greenwood, M., Flack, M., etc.: *see* Hill's Caisson Sickness.
[35]Haldane, J. S.: Deep Water Diving, Committee of the Admiralty (British), *see* Hill's Caisson Sickness.
[36]Cotton, T. F., Rapport, and Lewis, T.: Heart, 1918.
[37]Hill, Leonard, and Macleod, J. J. R.: Jour. Physiol., 1908, xxxvii, 77.
[38]Patterson, S. W., Piper, H., and Starling, E. H.: Jour. Physiol., 1914, xlviii, 465.

PART V

DIGESTION

CHAPTER XLVIII

GENERAL PHYSIOLOGY OF THE DIGESTIVE GLANDS

The function of digestion is to bring the food into such a condition that it can be absorbed through the intestinal epithelium into the blood and lymph. Carbohydrates are broken down as far as monosaccharides; neutral fats are split into fatty acids and glycerine; and proteins are broken down into the amino acids. The agencies which effect these decompositions are the digestive enzymes, or ferments, contained in the various digestive fluids or juices. The digestive juices are produced by glands, which are most numerous in the upper levels of the gastro-intestinal tract, the lower levels having as their main function that of absorption of the digested products. In order that the masses of food may be kept in a state of proper consistency, and that they may move readily along the digestive canal, numerous mucous glands are also scattered along the whole extent of the canal. Some of the digestive glands, such as the main salivary glands, the pancreas, and the liver, discharge their secretions into the digestive canal by special ducts, whereas others, such as the isolated salivary gland follicles in the mouth, the gastric glands and the crypts of Lieberkühn in the intestine, do not have an anatomically distinct duct, but discharge their secretions directly into the digestive tube.

It will be convenient to consider, first of all, certain properties that are common to the digestive glands, and then, the conditions under which each gland functionates during digestion.

MICROSCOPIC CHANGES DURING ACTIVITY

Structurally the active part of the glands, represented by the acinus or tubule, is composed of a basement membrane lined internally with the secreting epithelium. Outside the basal membrane are the lymph spaces and blood capillaries. After the gland has been at rest, the cells become

filled with granules or small globules, which are often so numerous as almost entirely to obliterate the nucleus. When the gland becomes active, on the other hand, the granules or globules leave the cells, except for a few which remain toward the lumen border. (Figs. 143 and 144.)

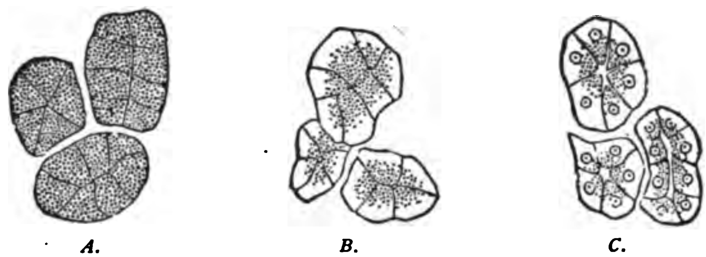


Fig. 143.—Cells of parotid gland showing zymogen granules: *A*, after prolonged rest; *B*, after a moderate secretion; *C*, after prolonged secretion. (From Langley.)

These observations indicate that the granular or globular material must represent part at least of the secretion of the glands. Sometimes, even before they are extruded, the granules become changed into some different material, as is indicated by the fact that they stain differently from

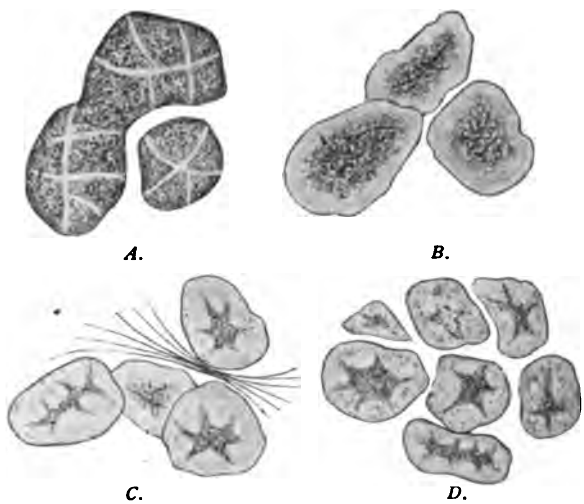


Fig. 144.—Parotid gland of rabbit in varying states of activity examined in fresh state. The upper left-hand acini are resting. The upper right-hand acini are from a gland stimulated to activity by injecting pilocarpine, and the two lower acini from one after stimulation of its sympathetic nerve. (After Langley.)

those of the resting gland. It must not be thought, however, that an extrusion of granules *necessarily* accompanies secretory activity, for under certain conditions a copious secretion of water and inorganic salts, as well as a certain amount of organic material, may be produced with-

out any change in the arrangement of the granules. In such cases it has been observed, as in the pancreas, that fine channels develop in the protoplasm of the cell (see page 429).

From this histologic evidence it would appear that the gland cell during rest is endowed with the property of building up out of the protoplasm, as granules or globules, the material which is to serve as one of the main organic constituents of the secretion. It is commonly believed that this is the precursor of the active ferment of the secretion; hence its name, *zymogen.* It has been shown that the process of separation of the zymogen granules starts around the nucleus with the production of a basophile substance, which in hardened specimens sometimes takes the form of filaments. From this basophilic *ergastoplasm,* as it is called, the granules are gradually formed, and then for some time continue to undergo slight further changes, as is evidenced by the fact that the staining reaction of those near the base of the cells differs from that of those at the free margin. When the gland cell is excited to secrete, the granules before being extruded, as noted above, often undergo a definite change, becoming swollen and more globular in shape.

MECHANISM OF SECRETION

These histologic studies merely tell us that active changes, associated with the production and liberation of certain of the constituents of its secretion, are occurring in the gland cell, but they throw no light on the mechanism whereby the gland cells secrete *water and inorganic salts.* This may be dependent, to a certain extent at least, on differences in osmotic pressure (see page 11). A possible explanation of the flow of water is as follows: If a watery solution of some osmotically active substance is put in a tube, which is closed at one end by a membrane impermeable to this substance and at the other by one permeable to it, and the tube immersed in water, a continuous current will be found to issue from the permeable end so long as there remains any osmotically active substance in the tube. If we assume, then, that the membranes at the two ends of the secreting cell are of such a nature that the one next the basement membrane is impermeable to some osmotically active substance manufactured by the cell, and the other toward the lumen is permeable, it will be clear that, so long as this substance exists in the cell, it will attract water from the blood, and the water together with the osmotically active substance will be discharged into the lumen.

It is possible that when anything excites the cell to secretory activity, such as a nerve impulse or hormone, it does so by causing a change in

the permeability of the lumen border of the cell. This change in permeability may be dependent upon alterations in surface tension brought about by the migration of electrolytes to the border. That such a migration of electrolytes does actually occur has been demonstrated by A. B. Macallum[8] who developed a microchemical test for potassium, and was able to show that this electrolyte accumulates at the lumen border of the cell during secretory activity. Potassium may be taken as a prototype of electrolytes in general. Support is given to this view by the fact that potassium always accumulates at the border of the cell through which the secretion takes place. In the epithelium of the small intestine, for instance, where the current goes in the opposite direction to that in gland cells, the accumulation of potassium occurs at the portion of the cell next the basement membrane.

Another possibility is that, when the gland becomes more active, the molecules present in the cell become broken down into smaller molecules and so raise the osmotic pressure of the cell content, with the result that water is attracted from the blood.

When the gland is excited so that the *zymogen granules,* as well as water and salts, are secreted, the primary change appears to involve the granules only. Those near the lumen swell up by absorbing water, and become converted into spheres in which salts are dissolved in smaller proportions than exist in the lymph bathing the cells. These swollen structures are then ruptured at the periphery of the cell and discharged into the lumen. This discharge of a fluid containing fewer saline constituents than the cell or surrounding blood plasma brings about increased concentration in the remaining parts of the cell, a process which possibly is assisted by a breaking up of molecules in the protoplasm itself, and which causes an increase in osmotic pressure with a consequent flow of water from the lymph to the cells and therefore from the blood to the lymph.

OTHER CHANGES DURING ACTIVITY

Whatever may be the nature of the physiological changes that are responsible for the secretory activity of the cell, the fact stands out prominently that a considerable expenditure of energy is entailed. This is indicated by the fact that considerably larger quantities of oxygen are taken up by the gland when it is in an active state than when at rest. Thus, the *oxygen consumption* of the resting submaxillary gland of the cat may be increased five times during active secretion. On account of this increased oxygen consumption it is not surprising that it should be found that the secretory activity of the cell is greatly impaired by a deficiency in oxygen.

These active processes occurring in the gland when it is excited to secrete are associated with changes in electric reaction and in the volume of the gland. The *electric changes* have been most extensively studied in connection with the salivary gland. Cannon and Cattel,[6] by connecting a galvanometer with nonpolarizable electrodes, one placed on the gland and the other on neighboring connective tissue, were able to show that with each period of active secretion a current of action was set up. This was first discovered by Rose Bradford and Bayliss, and has been carefully studied by Gesell.[6a] That the electric current is definitely associated with the secretion of saliva and is not caused by the vascular changes which usually accompany this act was shown by its occurrence when the blood supply was shut off from the gland, and by its absence when there was no secretion even though the vascular changes were brought about; neither is the electric change due to the movement of fluid along the duct, as evidenced by its persistence after ligation of the duct.

With regard to change *in volume,* it might be expected, on account of the greater vascularity of the gland accompanying activity, that this would increase. On the contrary, however, it has been shown to decrease, because of the large quantity of fluid secreted from the gland cells.

The action of two *drugs* on the gland cells is of considerable physiologic importance: that of atropine, which paralyzes the secretion, and that of pilocarpine, which stimulates it. We shall see later how this information may be used in working out the exact mechanism of the different glands.

Important observations concerning *the relationship of glandular activity to the blood supply* have been made by experiments in which glands were artificially perfused outside the body. When the submaxillary gland of the dog is perfused with oxygenated Ringer's solution, stimulation of its nerve supply does not produce the usual secretion, but if the Ringer's solution is mixed with blood plasma, the nerve stimulation has its usual effect for a short time. Although no secretion occurs when oxygenated Ringer's solution is perfused alone, the usual vascular changes still occur in the gland. The results seem to indicate that the presence of some constituent of the blood plasma is essential for the change in the permeability of the cell wall necessary for the process of secretion. Similar results have been obtained during artificial perfusion of the pancreas when secretin was used as the stimulus.

CONTROL OF GLANDULAR ACTIVITY

Having outlined the general nature of the changes occurring in gland cells during their activity, we may now proceed to study the nature of

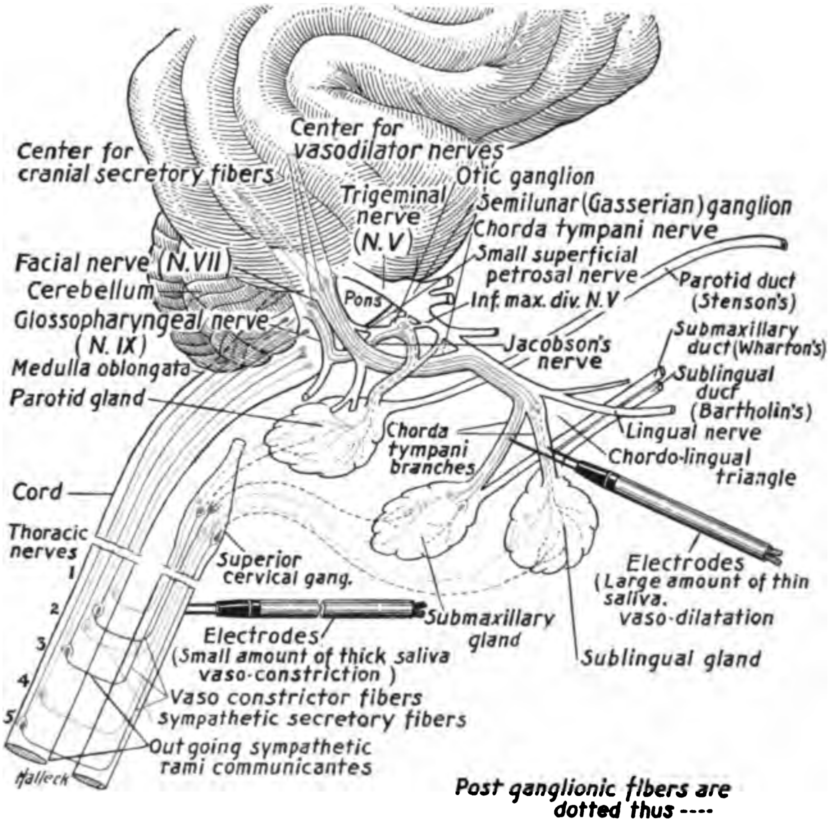


Fig. 145. —Diagrammatic representation of the innervation of the salivary glands in the dog. (From Jackson.)

the process by which this glandular activity is controlled. Two mechanisms of control are known: (1) by the nervous system, and (2) by means of hormones.

Nervous Control.—Control through the nervous system is most marked —indeed it may be the only means of control—in glands which have to produce their secretion promptly, whereas hormone control predominates in those in which prompt changes in secretory activity are not required. Thus, nervous control alone is present in the salivary glands, whereas hormone control is predominant in the pancreas, intestinal glands and liver. The gastric glands are partly under nervous control, and partly under hormone control. It should be pointed out here that the glands of the body other than the digestive glands are also subject to nervous or hormone control according to the promptness with which they are required to secrete. The lachrymal and sweat glands, and the venom glands of reptiles, for example, are practically entirely under nervous control, whereas most of the ductless glands, with the exception of the adrenals, are mainly under the influence of hormones.

The exact nature of the nervous control of glandular function has, therefore, been most extensively studied in the salivary glands, and that of the hormonic in the pancreas. With regard to the salivary glands, the following points are of importance: Their nerve supply comes from two sources: the cerebral autonomic, and the sympathetic autonomic (see page 877). These two nerve supplies have usually an opposite influence on the secretory activity of the glands, and very frequently also on the vascular changes that accompany secretory activity.

On account of its ready accessibility, the submaxillary gland in the dog and cat has been most thoroughly investigated. The cerebral autonomic nerve in this case is represented by the chorda tympani, and the sympathetic autonomic by postganglionic fibers that run from the superior cervical ganglion to the gland along its blood vessels (Fig. 145). After tying a cannula into the duct of the gland, it will be found *in the dog* that stimulation of the chorda tympani produces an immediate and abundant secretion of thin watery saliva accompanied by a marked dilatation of the blood vessels of the gland.

That this secretion is not dependent on the vasodilatation is easily shown by repeating the experiment after administering a sufficient dose of atropine to paralyze the secreting cells. Stimulation of the nerve then produces a vasodilatation but no secretion. The same conclusion is arrived at by an experiment of an entirely different nature; namely, by observing the pressure produced in the duct when the chorda tympani is stimulated. This pressure rises considerably above that in the arteries, so that no such physical process as mere diffusion can be held accountable

for the secretion, and therefore vasodilatation alone can not be responsible for it. If the sympathetic nerve supply is stimulated, a very scanty, thick secretion takes place accompanied by vasoconstriction.

Repetition of these experiments *in the cat* yields different results, particularly with regard to the influence of the sympathetic, a copious secretion being produced by such stimulation. The histologic changes produced in the gland cells are marked after sympathetic stimulation, but very slight, if present at all, after chorda stimulation.

The outstanding conclusion which may be drawn from these results is that two kinds of secretory activity are mediated through the nerves; one causing a thin watery secretion, containing only a small percentage of organic matter, and the other, a thick viscid secretion with a large amount of organic material. To explain these differences the hypothesis has been advanced that there are really two kinds of secretory fibers, called *secretory* and *trophic*, the former having to do with the secretion of water and inorganic salts, and the latter with the secretion of organic matter; i. e., with the extrusion of the zymogen granules. Certain authors (Langley) believe that such an hypothesis is unnecessary, and that the different results are dependent upon the concomitant changes in the blood supply produced by stimulating one or other nerve.

That there are really different kinds of true secretory fibers is, however, evident from the following experiment. If the duct of the gland is made to open through a fistula in the cheek, secretion of saliva through the fistula can be induced by placing various substances in the mouth, such as meat powder or hydrochloric acid. If the experiment is performed in such a way that the bloodflow through the gland can be observed, it will be found that the saliva produced by the stimulation with the meat powder contains a very much higher percentage of organic material than does that produced when hydrochloric acid is the stimulant, whereas the vascular changes in the gland and the inorganic constituents of the saliva are the same in both cases. Since stimulation of the chorda tympani causes the secretion of a watery saliva, while that caused by stimulation of the sympathetic is thick, it might be thought that the secretory fibers were contained in the former and the trophic fibers in the latter nerve; that this is not the case can be shown by a repetition of the above experiment in animals from which the superior cervical ganglion has been removed. The same results are obtained, indicating that the chorda tympani contains both secretory and trophic fibers.

CHAPTER XLIX

PHYSIOLOGY OF THE DIGESTIVE GLANDS (Cont'd)

THE HORMONE CONTROL

This is exhibited best in the case of the pancreas. The crucial experiment demonstrating that this gland is not primarily dependent upon nervous impulses for the control of its activity was performed by Bayliss and Starling.[2] Starting with the well-known fact that the application of weak acid to the duodenal mucous membrane excites secretion of pancreatic juice, these workers carefully severed all the nerve connections of a portion of the duodenum, and found on again applying acid to the mucous membrane that the secretion persisted. To explain this result they postulated that the acid must cause some substance to be liberated into the blood stream, which carries it to the pancreas, the cells of which it then excites to activity. To test this hypothesis they scraped off the mucous membrane of the duodenum and ground it in a mortar with weak hydrochloric acid (0.6 per cent), and, after boiling the solution so as to remove the protein and nearly neutralizing it, they obtained a fluid which, when injected intravenously, immediately caused a copious secretion of pancreatic juice.

Accompanying the secretion, however, a marked fall in arterial blood pressure was observed, making it possible that the secretion might have been due to a vasodilatation occurring in the pancreatic blood vessels. To eliminate this possibility they prepared an extract that was free of the depressor substances by extracting intestinal epithelium without any of the submucous tissue. The resulting extract had merely the secretory effect and produced no fall in blood pressure. This secretagoguary substance they named *secretin.*

Further evidence that the action of secretin is independent of the depressor substances has been obtained by taking advantage of the fact that the depressor substance is more soluble in alcohol than the secretin. If an acid decoction of duodenal mucous membrane is poured into absolute alcohol, a precipitate is formed. If this precipitate is redissolved in water and reprecipitated several times by absolute alcohol, then after drying a white powder is obtained, which is easily soluble in water. The resulting solution injected intravenously has a powerful secretory action, but produces no effect on blood pressure. The concentrated alcoholic

liquor, on the other hand, when similarly injected produces a marked fall in blood pressure. It is believed that this effect is due to the action of β-imidazolylethylamine. A very strong preparation of secretin can also be prepared by the method of Dale and Laidlaw,[7] which depends on precipitation by mercuric chloride.

Secretin does not exist preformed in the epithelial cells, as is shown by the fact that an extract, made with neutral saline solution, does not as a rule, have any secretory action when injected intravenously. Sometimes a slight secretion may be produced, but this is probably to be explained by the fact that some secretin remains behind in the cells as a result of a preceding phase of activity. If, on the other hand, the above neutral or slightly alkaline opalescent solution of the mucous membrane is boiled with acid, secretin may become developed in it. The interpretation put upon these results is that a substance, called *prosecretin*, exists in the epithelial cells, and that this becomes converted into secretin by the action of acid on the cells. The secretin thus produced is then taken up by the blood, none of it passing into the intestinal canal, because the free borders of the cells are impervious to secretin. That this is actually the case has been shown by finding that the introduction of neutralized secretin solution into the duodenum, or other parts of the small intestine, does not cause a secretion of pancreatic juice.

We know practically nothing concerning the *chemical nature of secretin*. Being soluble in about 90 per cent alcohol and in fairly weak acids, it can not belong to any of the better known groups of proteins. As it is readily diffusible through parchment membrane, it can not be of very complex structure, and as it withstands heat, it can not be an enzyme. It rapidly deteriorates in strength in the presence of alkalies.

Any acid when applied to the mucous membrane is capable of producing secretin, and so are certain other substances, such as mustard oil. Watery solutions of saccharose or urea, when rubbed up with the duodenal mucosa in a mortar, produce secretin solutions of varying activity, but they do not in the living animal excite pancreatic secretion when applied to the duodenum. Secretin is very susceptible to destruction by such digestive enzymes as those present in the pancreatic, gastric, and intestinal juices. That secretin is present in the blood when acid is in contact with the duodenal mucosa has been shown by the fact that injection into a normal dog of blood from one in which secretin formation is going on (as a result of acid in the duodenum), excites pancreatic secretion.

The pancreatic juice produced by the injection of secretin, like that which is produced under normal conditions, does not contain any active trypsin, but instead contains its precursor, trypsinogen. This becomes converted into trypsin in the intestine, being activated by contact with

enterokinase, an enzyme present in the intestinal juice. By such a mechanism the mucosa of the pancreatic duct is protected against autodigestion by trypsin.

NERVOUS CONTROL OF PANCREAS

Prior to the discovery of secretin, Pavlov[1] and his pupils had published numerous experiments purporting to show that the secretion of pancreatic

Fig. 146.—Pancreatic acini stained with hematoxylin. The acini at the top and to the left of the figure are from a resting gland, those to the right being from one that had been secreting for over three hours as a result of acid in the duodenum. The lowermost figure is from a gland the vagus nerve supply of which had been stimulated off and on for several hours. Note that the zymogen granules are extruded only after vagus activity but not after secretin activity. (From Babkin, Rubaschkin and Ssawitsch.)

juice is controlled through the vagus nerve. The amount of secretion produced by nervous stimulation was, however, never found to be so large as that produced by secretin, and for several years after the discovery of

the latter hormone, much doubt existed as to the correctness of Pavlov's claim. As in many other fields of physiologic science, investigators attempted to show that one or the other mechanism obtained, and they were not inclined to consider the possibility that both mechanisms might exist side by side. That such is the case, however, is clear from the most recent work, in which it has been found that if proper precautions are taken, repeated stimulation of the vagus nerve does call forth a secretion of pancreatic juice which, besides being less copious than that following

I.

II.

III.

Fig. 147.—Three preparations of pancreatic acini stained by eosin orange toluidin blue. The acini of Fig. I were from a gland after vagus stimulation, and it is noted that besides free extrusion of the granules, globules staining with orange (and appearing in deep black in the photograph) have formed and may be present in the ductules. Some of the globules, however, change in their staining properties, becoming light red (dark gray in photograph). The acini in II and III were from glands excited by secretin. No globules appear; the granules remain, and fine canaliculi appear in the clear protoplasm. (From Babkin, Rubaschkin and Ssawitsch.)

secretin injection, differs from it in the important fact that it contains not trypsinogen but active trypsin. Since the normal pancreatic juice contains trypsinogen, this last mentioned fact would appear to indicate that vagus control of the normal secretion can not be an important affair. The vagus secretion of pancreatic juice is, moreover, paralyzed by atropine, which has no action on the secretin mechanism (cf. Bayliss).

The copious secretion of pancreatic juice produced by secretin, on the one hand, and the scanty, thick secretion produced by vagus stimulation, on the other, calls to mind similar differences observed in the secretion of saliva as the result of chorda-tympani or sympathetic stimulation. It will be remembered that from these latter results it was concluded that there must be secretory and trophic fibers concerned in the control of the activities of gland cells. Interesting corroboration of this conclusion has recently been obtained by *histologic examination of the pancreas following secretin or vagus activity*. After the repeated injection of secretin, it is difficult to observe any signs of fatigue in the cells; the zymogen granules remain practically as numerous as in a resting gland, but in the clear protoplasm of the outer third of the cell, it is said that fine channels of fluid can be seen. Through these channels water is believed to pass from the blood towards the lumen and in its course to carry with it some of the zymogen granules, without, however, changing them. Thus, when the gland cells are stained with eosin and orange, after secretin activity some of the zymogen granules can occasionally be seen in the lumen of the acini stained with eosin like those in the cell itself. After vagus stimulation the appearances are different; not only are the granules more freely extruded from the cells, but they undergo a preliminary change; they lose the property of staining with eosin and become stained with the orange, at the same time increasing in size so as to form vacuoles. These vacuoles may wander into the ductules, and when they are present here they are stained by orange (Figs. 146 and 147) (Babkin, etc.[7a]).

Why there should be both a nervous and a hormone control of the pancreatic secretion is not clear. This gland, unlike the gastric and salivary glands, is not called upon to become active all of a sudden, and it is difficult to see what could serve as the normal stimulus operating through the nervous pathway. Taking it all in all, it is probably safe to conclude that the nervous mechanism is relatively unimportant, and that under normal conditions it seldom if ever is called into operation. Corroboration for this view is afforded by the fact, above mentioned, that the pancreatic juice produced by vagus stimulation contains active trypsin, which is not the case with normal pancreatic juice.

CHAPTER L

PHYSIOLOGY OF THE DIGESTIVE GLANDS (Cont'd)

Up to the present we have been concerned with the physiologic activities of digestive glands in general, but now we must study each of them separately in order to find out the conditions under which they become stimulated to activity in the normal process of digestion. The secretion of each gland has a definite role assigned to it in the complex and lengthy process of digestion. It takes up its work where the preceding secretion left off; e. g., the pepsin of gastric juice digests protein so far as proteoses and peptone; the trypsin of pancreatic juice then attacks the proteoses and peptone, and the resulting lower degradation products are finally attacked by the erepsin of the intestinal juice. The secretions of the various glands are, therefore, required in a certain definite order—they are correlated; and we must now give some attention to the precise condition upon which the activity and correlation depend.

THE NORMAL CONDITIONS UNDER WHICH THE GLANDS BECOME STIMULATED TO INCREASED ACTIVITY

To make possible such observations on the normal activities of the glands, a preliminary operation has to be performed so as to bring the duct of the gland to the surface of the body and permit of the observation of its secretory activity after the animal has recovered from the immediate effects of the operation. We owe to Pavlov[1] the surgical technic by which these conditions can be fulfilled. The general principle of the operation, in the case of glands provided with ducts, consists in making a circular cut through the mucous membrane surrounding the opening of the duct and then, after dissecting the duct free, stitching the edges of the cut to the skin wound. Healing then takes place without the formation in the duct of any stricture due to the cicatricial tissue. After the wound has healed, the secretion can readily be collected in a receiver attached over the duct fistula, the animal being in every other way in a perfectly normal condition. In the case of glands not provided with a duct, other methods must be adopted to collect the secretions. These will be described elsewhere.

THE NORMAL SECRETION OF SALIVA

The duct fistula can in this case be made either for the submaxillary gland, representing a mucous gland, or for the parotid, representing a serous gland. Under ordinary conditions there is very little secretion from either duct. When secretion occurs, it is, of course, caused by influences acting on a nerve center or centers in the medulla oblongata, the exact location of which for the different glands has been worked out in recent years by Miller.[9] The impulses acting on these centers may be transmitted along afferent nerves coming from the mucous membrane of the mouth, nares, etc., or by impulses which we may call psychic, transmitted from the higher nerve centers. The reflex secretions caused by impulses traveling by the afferent nerve from the mouth, etc., have been called *unconditioned,* and those from the higher nerve centers, *conditioned.* With regard to the former, there is considerable discrimination in the type of stimulus that will be effective. Thus, if the dog—for most of the experiments have been performed on this animal—is given meat, a secretion of thick, mucous saliva will be observed to occur (submaxillary gland). On the other hand, if the meat is dried and pulverized, the secretion which it calls forth will be very copious and watery (parotid gland). There is, then, an obvious association between the nature of the secretion and the function it will be called upon to perform when it becomes mixed with the food. The mucous secretion called forth by meat will serve to lubricate the bolus of food and thus facilitate its swallowing, whereas the thin watery secretion produced by the dry powder will have the effect of washing the powder from the mouth.

It is evident that the mechanical condition of the food partly determines its exciting quality. Mechanical stimulation of the mucosa alone is, however, not an adequate stimulus, for if pebbles are placed in the mouth, little secretion occurs, but if sand is placed in the mouth, secretion immediately becomes copious. The nerve endings also respond to chemical stimuli. Thus, weak acid causes a copious secretion, while alkali has no effect; disagreeable, nauseous substances also excite secretion. The above differences in the response of the glands according to the mechanical condition of the food has been observed only in the case of the parotid gland, increase in the submaxillary secretion being obtained only when actual foodstuffs are placed in the mouth.

The investigations that have been made on the conditions of psychic secretion of saliva are still more interesting and important. Their importance depends not so much on the information they give us concerning the secretion of saliva as such, as on the methods they furnish us for investigating the various conditions that affect the psychic processes

associated with the taking of food. It is from the psychic rather than from the physiologic standpoint, therefore, that these observations are of importance, for they permit us, by objective methods, to study on dumb animals problems that would otherwise be beyond our powers of investigation. Many of the results, with their bearing on the functions of the higher nerve centers, have been discussed elsewhere. Meanwhile, however, even at the risk of repetition it may not be out of place to cite a few of the most interesting experiments.

If we tease a hungry animal with food for which he has a great appetite, a copious secretion of saliva immediately occurs. If we go on teasing him without giving him food, and repeat this procedure on several succeeding days, it will be found that gradually he no longer responds to the teasing by increased salivation. Evidently, therefore, the reflex is conditioned upon the animal's afterward receiving the food.

The experiment may be performed in another way. If, for example, we offer the animal some food for which he has no appetite, no secretion of saliva will occur; but, if at the end of the process we give him appetizing food, it will be found after repeating this procedure on several successive days that the presentation of the unappetizing food calls forth a secretion. He has learned to associate the presentation of unappetizing food with the subsequent gratification of his appetite. The experiment can even be performed so that a definite interval of time elapses between the application of the stimulus and the salivation: if the animal is teased on successive days with food for which he has an appetite but is not given the food until after ten or twenty minutes, presentation of this food will come to be followed by salivation—not immediately, but after the exact interval of time that had been allowed to intervene in the training process. During this interval there must be an inhibition of psychic stimulation of the salivary centers by other nerve centers. It is of great interest that this inhibition may itself be inhibited by various forms of stimulation of the nervous system (see page 858).

THE SECRETION OF GASTRIC JUICE

Methods of Investigation

There being no common duct, the secretion of the gastric glands is a much more difficult problem to investigate than is that of glands which, like the salivary, are supplied with ducts. One of the most interesting chapters in the history of physiology concerns the methods which from time to time have been evolved for the collection of this juice and for studying the digestive processes in the stomach. Prominent among the problems confronting the earlier investigators was the question whether

the main function of the stomach is to crush or triturate the food or to act on it chemically. The great French scientist Reaumur and a little later the Italian Abbé Spallanzani (1729-1799) attacked this problem by methods that anticipate those of Rehfuss and Einhorn. Spallanzani ultimately devised the method of swallowing small perforated wooden tubes containing foodstuffs and covered by small linen bags. After the bags were passed per rectum, he found that considerable erosion or digestion of the food had occurred, but that the wooden tubes, however thin-walled they might be, were not crushed. In order to secure samples of the gastric juice free from food, the only method available to the older investigators consisted in swallowing sponges attached to threads, which after being for some time in the stomach were withdrawn and squeezed dry of juice.

The next great contribution came from this country, where, in 1833, Dr. Beaumont, while a surgeon in the service of the American troops located at Mackinaw, made observations on a Canadian voyageur by the name of Alexis St. Martin, who by the premature discharge of his gun had wounded himself in the stomach, the wound never healing but leaving a permanent gastric fistula. Beaumont arranged to keep Alexis St. Martin in his service for several years, during which time he made numerous observations on the process of digestion in the stomach—observations many of which are of great value even at the present day.

By none of these methods, however, could a sample of pure gastric juice be secured while the digestive process was actually in progress. To make the collection of such a sample possible, Heidenhain devised a method of isolating portions of the stomach wall as pouches opening through fistulæ on the abdominal wall. The results of Heidenhain's experiments are, however open to the objection that the secretion in the isolated pouches may not really correspond to that occurring in the main stomach, since the connections of the pouches with the central nervous system must have been severed. In order that these connections might remain as nearly intact as possible, the Russian physiologist, Pavlov,[1] devised an ingenious operation in which the pouch, or "miniature stomach," remains connected with the main stomach through a considerable width of mucous and submucous tissue. The essential nature of this operation will be evident from the accompanying diagram. (Fig. 148).

The most recent investigations have been made by Cannon[3] and by Carlson.[4] The former fed animals food impregnated with bismuth subnitrate, and then exposed the animal to the x-rays. A shadow is produced by the food mass in the stomach, and from the changes in the outline of this shadow facts have been collected, not only concerning the

movements of the viscus, but also concerning the rate of discharge of food into the intestine and therefore the duration of the gastric digestive process. Carlson's contribution has been rendered possible by his good fortune in having in his service a second Alexis St. Martin, a man with complete closure of the esophagus and a gastric fistula large enough to permit of direct inspection of the interior of the stomach. Seizing the opportunity thus presented, Carlson during the last four or five years has devoted his attention exclusively to a thorough investigation, not only of the movements of the stomach, but also of the rate of secretion of the gastric juice under different conditions. He has also, with praiseworthy enthusiasm and keen scientific spirit, extended his observations both on laboratory animals and on himself and his coworkers, so as not

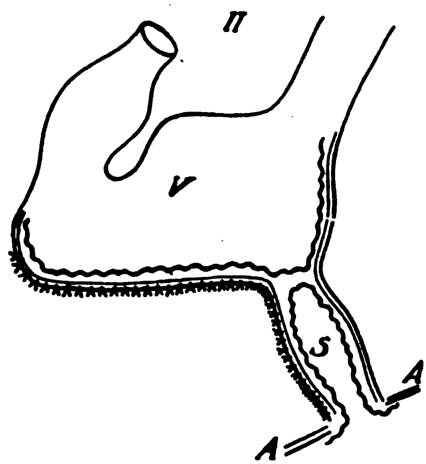


Fig. 148.—Diagram of stomach showing miniature stomach (*S*) separated from the main stomach (*V*) by a double layer of mucous membrane. *A.A.*, is the opening of the pouch on the abdominal wall. (Pavlov.)

to incur the error, which is all too frequently made, of confining the observations to one animal.

The Nervous Element in Gastric Secretion

The first stimulus to the secretion of gastric juice is nervous in origin, and is dependent on the gratification of the appetite and the pleasure of taking food. This fact, after having been suggested by observations made in the clinic, was first thoroughly investigated by Pavlov, who for this purpose observed the gastric secretion flowing either from a fistula of the stomach itself, or from a "miniature stomach," in dogs in which also an esophageal fistula had been established. When food was given by mouth to these animals, it was chewed and swallowed in the usual manner, but before reaching the stomach, it escaped through the esopha-

geal fistula. This experiment is known as that of "sham feeding." Within a few minutes after giving food the gastric juice was found to be secreted actively, and if the feeding process was kept up, which could be done almost indefinitely since the animal never became satisfied, the secretion continued to flow. Thus, in one instance Pavlov succeeded in collecting about 700 c.c. of gastric juice after sham feeding an animal for five or six hours in the manner above described.

After the stomach has emptied itself of the food taken with the previous meal, it is said by Pavlov to contain only a little alkaline mucus. The more recent work of Carlson, however, shows that this is not strictly the case, there being more or less of a continuous secretion of gastric juice in the entire absence of food. The amount varies from a few c.c. up to 60 c.c. per hour, more secretion being produced when it is collected every five or ten minutes than if it is collected every thirty or sixty, thus indicating that, ordinarily, some escapes through the pylorus into the duodenum. The secretion contains both pepsin and hydrochloric acid. As to the cause of this continuous secretion, little is known. It may be an example of the periodic activities of the digestive glands described by Boldyreff, or it may in part be due to a psychic stimulation dependent upon the thought of food. That the latter is probably not the cause, is indicated by the fact that, at least in Carlson's patient, the psychic juice could not be made to flow short of giving food.

The sham feeding causes stimulation of the gastric secretion through impulses transmitted to the stomach along the vagus nerves; for it has been found, in animals in which the vagus nerve has been cut, that the sham feeding no longer induces a secretion of gastric juice. The question therefore arises as to how the nerve center is stimulated. Three possible causes may be considered: (1) mechanical stimulation of the sensory nerves of the mouth; (2) chemical stimulation of the nerves; (3) the agreeable stimulation of the taste buds and olfactory endings concerned in the tasting of food. In investigating these possibilities, mechanical stimulation was readily ruled out by showing that mere taking of solid matter in the mouth did not excite any secretion, although it might cause a flow of saliva. Mere chemical stimulation could not be the cause, for no secretion was induced by placing substances such as acetic acid or mustard oil in the mouth. By exclusion, then, it would appear that the adequate stimulus must consist in the agreeable stimulation of the taste buds, etc.—that is to say, in the *gratification of appetite*.

Further justification for this conclusion was readily secured by noting that foodstuffs for which the animal had no particular desire or appetite failed to excite the secretion. Most dogs, for example, although they may take it, are not particularly fond of bread, and when fed with

it, these animals did not produce any appetite juice. In one animal that showed considerable liking for bread, active secretion occurred when he was fed with this foodstuff.

Pavlov further noted that usually it was not necessary actually to allow the animal to take the food into his mouth, but that mere teasing with savory food was sufficient to cause the secretion, and that in highly sensitive animals even the noises and other events usually associated with feeding time were sufficient to excite the secretion. In the case of a hungry animal, the mere approach of the attendant with food, or some other noise or action definitely associated with feeding time, was sufficient to excite the secretion. The appetite juice when started was found to persist for some time after the stimulus causing it had been removed.

Carlson has succeeded in confirming in man most of these observations. He noted, however, that the secretion produced by seeing or smelling or thinking of food is much less than would be expected from Pavlov's observations on dogs. Even when his subject was hungry, Carlson did not observe that the bringing of a tray of savory food into the room caused any secretion of gastric juice. It is, of course, to be expected that the quantity of the psychic secretion will not be the same in different individuals. It has been observed, for example, by Pavlov to vary considerably in the case of dogs, and it is very likely that it will vary still more in man, with his more highly complicated nervous system. In no case could Carlson observe any secretion of gastric juice produced by having his patient chew on indifferent substances, or by stimulating the nerve endings in the mouth by substances other than those directly related to food.

In man the rate of secretion is proportional to the palatability of the food, the smallest amount, during twenty minutes' mastication of palatable food, being 30 c.c. and the largest 150 c.c., in a series of 156 observations. A typical curve showing the amount of the secretion is given in Fig. 149. To construct this curve the gastric juice was collected during five-minute intervals while the man was chewing a meal of average composition and of his own choice. An interesting feature depicted on this curve is that the secretion rate was highest in the last five-minute period, this being the time during which the dessert was being taken, for which this man had a great relish. Quite clearly there was a direct relation between the rate of the secretion of the appetite juice and the palatability of the food. It will further be observed that it took only from fifteen to twenty minutes after discontinuing the chewing before the juice returned to its original level.

The practical application of these facts in connection with the hygiene

of diet and the feeding of patients during convalescence, is obviously very great. However perfect in other regards a diet may be, it will probably fail to be digested at the proper rate unless it is taken with relish. Frequent feeding with favorite morsels is more likely to be followed by thorough digestion and assimilation than occasional stuffing with larger amounts. We see too in these experiments an explanation of the well-established practice of starting a meal with something savory. A *hors d'oeuvre* is nothing more than a physiologic stimulant to appetite. It is also interesting from a practical standpoint to observe that with those who have a keen relish for sweetmeats the taking of dessert has a real physiologic significance, for, as in Carlson's patient, it stimulates toward the end of a meal a further secretion of the gastric

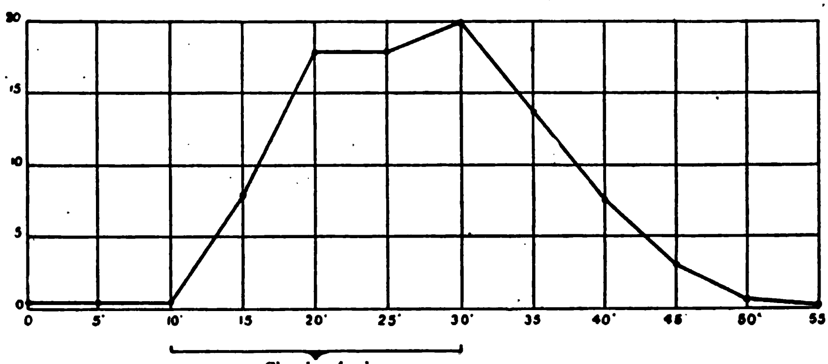


Fig. 149.—Typical curve of secretion of gastric juice collected at 5-minute intervals on mastication of palatable food for 20 minutes. The rise in secretion during the last 5 minutes of mastication is due to chewing the dessert (fruit) for which the person had great relish. (From Carlson.)

juice, and thus insures a more rapid digestion of the food. Good cooking, it should be remembered, is really the first stage in digestion, and it is the only stage over which we can exercise voluntary control.

The Hormone Element in Gastric Secretion

Although gastric digestion is initiated by the appetite juice, it is clear that this alone can not account for all the secretion that occurs during normal gastric digestion. After an ordinary meal gastric digestion lasts usually about four hours, whereas we have seen, particularly from Carlson's observations, that the appetite juice lasts only for some fifteen or twenty minutes after the exciting stimulus has been removed. The appetite juice, in other words, serves only to initiate the process of secretion, and the question arises, *What keeps up the secretion during the rest of gastric digestion?* The answer was furnished by Pavlov, who

observed animals in which not only a miniature stomach had been made, but a fistula into the main stomach as well. The behavior of the secretion of gastric juice as a whole could be followed by collecting that which was secreted in the miniature stomach, for it was shown, in control experiments, that this secretion runs strictly parallel with that in the main stomach, being quantitatively a definite fraction of it—according to the relative size of the miniature stomach—and qualitatively identical. The miniature stomach, in other words, mirrors the events of secretion in the main stomach.

It was observed that when the animal was allowed to take the food into the main stomach by the mouth and esophagus, the secretion from the miniature stomach continued to flow until the process of gastric digestion had been completed, a result which was quite different from that obtained after sham feeding. The only possible explanation for this result is that the food in the stomach sets up secretion as a result of *local stimulation.* To investigate the nature of this local stimulation, whether mechanical or chemical, food and other substances were placed in the main stomach through the gastric fistula without the animal's knowledge so as to avoid possible psychic stimulation, and the secretion observed from the miniature stomach. When the mucous membrane of the main stomach was stimulated mechanically, as by placing inert objects such as a piece of sponge or sand in the stomach, no secretion occurred. Evidently, therefore, the stimulus is dependent upon some chemical quality of the food.

By introducing various foods it was found that there is considerable difference in the degree to which they can excite the secretion. Water, egg white, bread and starch, were all found to have very little if any effect. On the other hand, when protein that had been partly digested by means of pepsin and hydrochloric acid was introduced into the stomach, it immediately called forth a secretion. The conclusion is that the partly digested products, even of insipid food, are capable of directly exciting the secretion. These include proteoses and peptones, and it was, therefore, of great interest to find that a solution of commercial peptone is also an effective stimulus. This is a result of deep significance, for it indicates that the food which has been partially digested by the appetite juice will serve as a stimulus to continued secretion.

The psychic juice has been aptly called the "ignition juice," because by producing partial digestion it serves to ignite the process of gastric secretion. Experimental evidence of its great importance in gastric digestion was secured by Pavlov in experiments in which he placed weighed quantities of meat attached to threads in the stomach through a gastric fistula, and after some time removed them and determined by

the difference in weights the extent to which they had become digested. It was found that when the appetite juice was excited by sham feeding at the same time that food was placed directly in the stomach, its digestion was much more rapid than in cases in which it was placed in the stomach without the animal's knowing, as when he was asleep.

Other foods having a direct stimulating effect on the gastric secretion are meat extracts and, to a certain extent, milk. This effect of meat extract is interesting in connection with the practice of taking soup as a first or early stage in dining. It not only excites the appetite juice, but also serves as a direct stimulus to the gastric secretion.

As to *the nature of the mechanism by which this direct secretion takes place,* it was shown by Popielski[10a] that the secretion still occurs after all the nerves proceeding to the stomach are cut. Evidently, therefore, it is independent of the extrinsic nerve supply of the viscus. As a result of his experiments Popielski concluded that the secretion must depend on a local reflex mediated through the nerve structures present in the walls of the stomach itself. Another explanation of the result has, however, in recent years been given more credence by the experiments of Bayliss and Starling on the influence of hormones on the secretion of pancreatic juice (cf. page 425). Edkins[10] suggested that a similar process in the stomach might account for the continued secretion of gastric juice. To test the possibility this investigator, after ligating the cardiac sphincter in anesthetized animals, inserted a tube into the pyloric end of the stomach, through which he placed in the stomach about 50 c.c. of physiologic saline. After this had been in the stomach for an hour, he found that no water was absorbed, and that if the fluid was removed after this time, it contained neither hydrochloric acid nor pepsin. On the other hand, if during the time the saline was in the stomach a decoction of the mucous membrane of the pyloric end, made either with peptone solution or with a solution of dextrine, was injected intravenously in small quantities every few minutes, it was found that the saline contained distinct quantities of hydrochloric acid and pepsin. Furthermore, it was found that, if the peptone solution or the dextrine solution alone was injected intravenously, there was no such evidence of gastric secretion. The conclusion which Edkins drew from his experiments is to the effect that the half-digested products of the earlier stages of gastric digestion act on the mucous membrane of the stomach so as to produce a hormone, which is then carried by the blood to the cells of the gastric glands, upon which, like secretin, it directly develops an exciting effect. This hormone has been called *gastrin.* These observations of Edkins have been confirmed, and they explain very simply how gastric secretion is maintained after the cessation of the secretion of the

appetite juice.[10] By such a mechanism gastric juice would continue to be secreted so long as any half-digested food remains in the stomach.

The action of gastrin is the first instance of a hormone control of the digestive glands. In the earlier stages of digestion, the secretion of saliva and appetite juice is mediated through the nervous system, because these juices must be produced promptly. In the later stages of gastric digestion, such promptitude in response on the part of the gland is no longer necessary, so that the slower, more continuous process of hormone control is sufficient.

Quantity of Gastric Juice Secreted

According to Carlson, the total amount of gastric juice secreted in man on an average meal composed of meat, bread, vegetables, coffee or

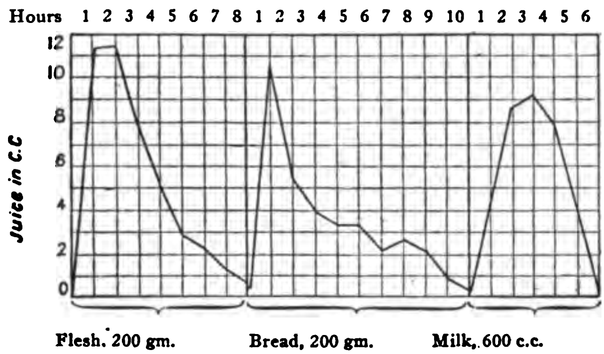


Fig. 150.—Cubic centimeters of gastric juice secreted after diets of meat, bread, and milk. (From Pavlov.)

milk, and dessert, amounts to about 700 c.c., being divided into 200 c.c. in the first hour, 150 in the second, and 350 c.c. during the third, fourth and fifth hours. These figures were estimated partly on the basis of observations made on the man with the gastric fistula, and partly from the data supplied by Pavlov's observations on dogs. Carlson believes that Pavlov overestimated the relative importance of the appetite juice in gastric digestion. He found, for example, that after division of both vagus nerves in dogs normal gastric digestion might be regained a few days after the operation, although, of course, under such circumstances no appetite juice could have been secreted. Moreover, he observed that cats when forcibly fed with unpalatable food may digest that food as rapidly as when they eat voluntarily. In support of his contention, Carlson states that he has frequently removed all of the appetite juice from his patient's stomach before the masticated meal was put into it without any evident interference with the digestive process.

Fat has a distinct inhibiting influence on the direct secretion of gas-

tric juice; cream takes considerably longer to be be digested than milk, and the presence of oil in the stomach delays the secretion of juice poured out on a subsequent meal of otherwise readily digestible food. By collecting all of the gastric juice from the miniature stomach after feeding by mouth with quantities of different protein-rich foods containing the same quantities of nitrogen, interesting observations have been recorded concerning the amount of juice secreted and its proteolytic power. The results of some of the experiments are shown in the accompanying curves (Figs. 150 and 151).

It will be seen that the most abundant secretion occurs with meat, that of milk being not only smaller but also slower in starting. The digestive power is greatest in the case of bread.

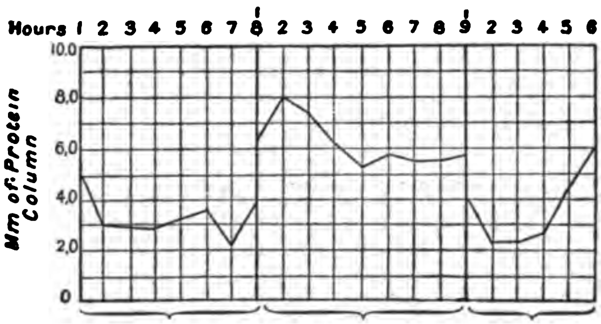


Fig. 151.—Digestive power of the juice, as measured by the length of the protein column digested in Mett's tubes, with diets of flesh, bread, and milk. (From Pavlov.)

THE INTESTINAL SECRETIONS

Pancreatic Juice

Regarding the natural secretion of pancreatic juice, little need be added to what has already been said (see page 425). The secretion begins when the chyme enters the duodenum, and attains its maximum when the outflow of this is greatest. By collecting the juice from a permanent fistula of the pancreatic duct, it has been found that the amount varies with different foods. When quantities of food containing equivalent amounts of nitrogen are fed, the greatest secretion is said to occur with bread and the least with milk. Such differences are probably dependent upon the amount of acid secreted in the stomach and passed on into the duodenum. It was thought at one time that, besides variation in quantity, the nature of the enzymes in the pancreatic juice might vary according to the kind of food. This, however, has been shown not to be the case.

Bile

The secretion of bile runs practically parallel with that of pancreatic juice. The liver is producing bile more or less continuously, since besides being a digestive fluid it is also an excretory product. The bile produced between the periods of digestion is mainly stored in the gall bladder. When the acid chyme comes in contact with the duodenal mucous membrane, it excites afferent nerve endings that cause a reflex contraction of the gall bladder, and this expresses some of the bile into the duodenum. The secretin, which the acid at the same time produces, besides affecting the pancreas, acts on the liver cells, stimulating them to the increased secretion of bile. Thus, by a nervous reflex operating on the gall bladder and later by a hormone mechanism operating on the liver cell, the increased secretion of bile is insured throughout digestion. Of the bile discharged into the intestine, a certain proportion of the bile salts is reabsorbed into the portal blood. When these arrive at the liver they also excite secretion of bile, thus assisting secretin in maintaining the secretion throughout the process of intestinal digestion.

Fig. 152.—Loop of intestine after tying off the portions, cutting the nerves running to the middle portion, and returning the loop to the abdomen for some time. (From Jackson.)

Intestinal Juice

The secretion of intestinal juice, or succus entericus, can obviously be studied only after isolating portions of the intestine and connecting them with fistulæ of the abdominal walls. It appears here again that both a nervous and a hormone mechanism exist. Mechanical stimulation of the intestinal mucous membrane causes an immediate outflow of intestinal juice, the purpose of which under normal conditions is evidently to assist in moving forward the bowel contents. This mechanically excited juice does not contain any enterokinase and only small amounts of the other enzymes. Further evidence for nervous control of the secretion of intestinal juice has been obtained by isolating three pouches of intestine be-

tween ligatures, and then denervating the central pouch by carefully cutting all the nerves without wounding the blood vessels. On returning the pouches to the abdomen and leaving them several hours, it has been found that the middle pouch becomes distended with secretion, whereas the two end pouches remain empty (Fig. 152). If the pouches are left for several days in the abdomen, however, the secretion from the denervated portion disappears again. The explanation of the result is possibly that the nerves under ordinary conditions convey impulses to the intestinal glands, which tonically inhibit their activity.

The existence of hormone control is evidenced by the fact that no enterokinase is present in the intestinal juice unless pancreatic juice is placed in contact with the mucous membrane. Injection of pancreatic juice into the blood, however, does not cause any secretion of intestinal juice; whereas the injection of secretin has such an effect.

CHAPTER LI

THE MECHANISMS OF DIGESTION

MASTICATION, DEGLUTITION, VOMITING

Mastication

By the movements of the lower jaw on the upper, the two rows of teeth come together so as to serve for biting or crushing the food. The resulting comminution of the food forms the first step in digestion. The up and down motion of the lower jaw results in biting by the incisors, and after the mouthful has been taken, the side to side movements enable the grinding teeth to crush and break it up into fragments of the proper size for swallowing. The most suitable size of the mouthful is about 5 c.c., but this varies greatly with habit. After mastication, the mass weighs from 3.2 to 6.5 gm., about one-fourth of this weight being due to saliva. The food is now a semifluid mush containing particles which are usually less than 2 mm. in diameter. Some, however, may measure 7 or even 12 mm.

Determination of the proper degree of fineness of the food is a function of the tongue, gums, and cheeks, for which purpose the mucous membrane covering them is supplied with very sensitive touch nerve endings (see page 794). The sensitiveness of the tongue, etc., in this regard explains why an object which can scarcely be felt by the fingers seems to be quite large in the mouth. If some particles of food that are too large for swallowing happen to be carried backward in the mouth, the tongue returns them for further mastication.

The saliva assists in mastication in several ways: (1) by dissolving some of the food constituents; (2) by partly digesting some of the starch; (3) by softening the mass of food so that it is more readily crushed; (4) by covering the bolus with mucus so as to make it more readily transferable from place to place. The secretion of saliva is therefore stimulated by the chewing movements, and its composition varies according to the nature of the food (page 431). In some animals, such as the cat and dog, mastication is unimportant, coating of the food with saliva being the only change which it undergoes in the mouth. In man the ability thus to bolt the food can readily be acquired, not, however, without some detriment to the efficiency of digestion as a whole. Soft

starchy food is little chewed, the length of time required for the mastication of other foods depending mainly on their nature, but also to a certain degree on the appetite and on the size of the mouthful.

It can not be too strongly insisted upon that the act of mastication is of far more importance than merely to break up and prepare the food for swallowing. It causes the food to be moved about in the mouth so as to develop its full effect on the taste buds; the crushing also releases odors which stimulate the olfactory epithelium. On these stimuli depend the satisfaction and pleasure of eating, which in turn initiate the process of gastric digestion (see page 435).

The benefit to digestion as a whole of a large secretion of saliva, brought about by persistent chewing, has been assumed by some to be much greater than it really is, and there has existed, and indeed may still exist, a school of faddists who, by deliberately chewing far beyond the necessary time, imagine themselves to thrive better on less food than those who occupy their time with more profitable pursuits.

Deglutition or Swallowing

After being masticated the food is rolled up into a bolus by the action of the tongue against the palate, and after being lubricated by saliva is moved, by elevation of the front of the tongue, towards the back of the mouth. This constitutes the *first stage* of swallowing, and is, so far, a voluntary act. About this time a slight inspiratory contraction of the diaphragm occurs—the so-called respiration of swallowing—and the mylohyoid quickly contracts, with the consequence that the bolus passes between the pillars of the fauces. This marks the beginning of the *second stage,* the first event of which is that the bolus, by stimulating sensory nerve endings, acts on nerve centers situated in the medulla oblongata so as to cause a coordinated series of movements of the muscles of the pharynx and larynx and an inhibition for a moment of the respiratory center (page 332).

The movements alter the shape of the pharynx and of the various openings into it in such a manner as to compel the bolus of food to pass into the esophagus (see Fig. 153): thus, (1) the soft palate becomes elevated and the posterior wall of the pharynx bulges forward so as to shut off the posterior nares, (2) the posterior pillars of the fauces approximate so as to shut off the mouth cavity, and (3) in about a tenth of a second after the mylohyoid has contracted, the larynx is pulled upwards and forwards under the root of the tongue, which by being drawn backwards becomes banked up over the laryngeal opening. This pulling up of the larynx brings the opening into it near to the lower half of the dorsal side of the epiglottis, but the upper half of this struc-

ture projects beyond and serves as a ledge to guide the bolus safely past this critical part of its course. (4) As a further safeguard against any entry of food into the air passages, the laryngeal opening is narrowed by approximation of the true and the false vocal cords.

So far the force which propels the bolus is mainly the contraction of the mylohyoid, assisted by the movements of the root of the tongue. When it has reached the lower end of the pharynx, however, the bolus readily falls into the esophagus, which has become dilated on account of a reflex inhibition of the constrictor muscles of its upper end. This so-called second stage of swallowing is, therefore, a complex coordinated movement initiated by afferent stimuli and involving reciprocal action of various groups of muscles: inhibition of the respiratory muscles and

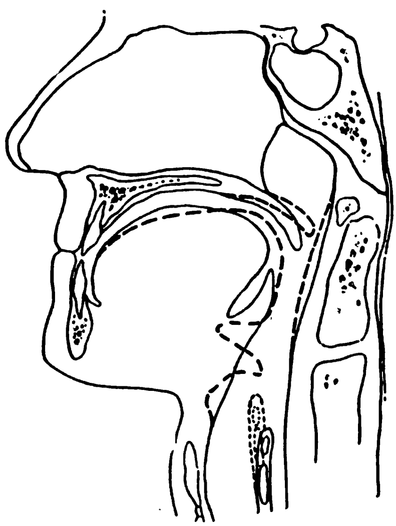


Fig. 153.—The changes which take place in the position of the root of the tongue, the soft palate, the epiglottis and the larynx during the second stage of swallowing. The thick dotted line indicates the position during swallowing.

of those that constrict the esophagus, and stimulation of those that elevate the palate, the root of the tongue, and the larynx. It is purely an involuntary process.

The *third stage of deglutition* consists in the passage of the swallowed food along the esophagus. The mechanism by which this is done depends very much on the physical consistence of the food. A *solid bolus* that more or less fills the esophagus excites a typical peristaltic wave, which is characterized by a dilatation of the esophagus immediately in front of and a constriction over and behind the bolus. This wave travels down the esophagus in man at such a rate that it reaches the cardiac sphincter in about five or six seconds. On arriving here the cardiac

sphincter, ordinarily contracted, relaxes for a moment so that the bolus passes into the stomach. In many animals, including man and the cat, the peristaltic wave travels much more rapidly in the upper part of the esophagus than lower down because of differences in the nature of the muscular coat, this being of the striated variety above, and of the non-striated below. The purpose of more rapid movement in the upper part is no doubt that the bolus may be hurried past the regions where, by distending the esophagus, it might interfere with the function of neighboring structures, such as the heart. In other animals, as the dog, the muscular fiber is striated all along the esophagus, and the bolus of food correspondingly travels at a uniform, quick rate all the way. It takes only about four seconds for the bolus to reach the stomach in the dog.

The peristaltic wave of the upper part of the esophagus in the cat and presumably in man, unlike that of the intestines (see page 466), is transmitted by the esophageal branches of the vagus nerves. If these are severed, but the muscular coats left intact, the esophagus becomes dilated above the level of the section and contracted below, and no peristaltic wave can pass along it; on the other hand, the muscular coat may be severed (by crushing, etc.) but the peristaltic wave will continue to travel, provided no damage has been done to the nerves.

In the lower part of the esophagus, however, the wave of peristalsis, like that of the intestines, travels independently of extrinsic nerves. This has been observed in animals in which all of the extrinsic nerves have been cut some time previously. This difference between the upper and the lower portions is associated with the difference in the nature of the muscular fibers above noted (Meltzer).[11]

The propagation of the wave by the nerves in the upper part of the esophagus indicates that the second stage and the first part of the third stage of deglutition must be rehearsed, as it were, in the medullary centers from which arise the nerve fibers to the pharynx and the upper levels of the esophagus. It is thought that the discharges from these local centers are controlled by a higher *swallowing center* situated in the medulla just above that of respiration, the afferent stimuli to which proceed from the pharynx by the fifth, superior laryngeal, and vagus nerves. The exact location of the sensory areas whose stimulation is most effective in initiating the swallowing reflex varies considerably in different animals. In man it is probably at the entrance to the pharynx; in the dog it is on the posterior wall. A foreign body placed directly in the upper portion of the esophagus of man has been observed to remain stationary until the individual made a swallowing movement. The afferent fibers in the glossopharyngeal nerve exercise a powerful *inhibitory* influence on the deglutition center as well as on that of respira-

tion. Thus, if swallowing movements are excited by stimulating the central end of the superior laryngeal nerve, they can be instantly inhibited by simultaneously stimulating the glossopharyngeal, and the respiratory movements stop in whatever position they may have been at the time. When the glossopharyngeal nerves are cut, the esophagus enters into a condition of tonic contraction, which may last a day or so. This shows that the inhibitory impulses are tonic in nature.

This inhibition of the esophagus is indeed a most important part of the process when *liquid or semiliquid food* is swallowed. By the contraction of the mylohyoid muscle, fluids are quickly shot down the distended esophagus, at the lower end of which, on account of the closure of the cardiac sphincter, they accumulate until the arrival of the peristaltic wave which has meanwhile been set up by stimulation of the pharynx. If the swallowing is immediately repeated, as is usually the case in drinking, the esophagus remains dilated because peristalsis is inhibited, and the fluid lies outside the cardiac orifice until the last mouthful has been taken.

The Cardiac Sphincter

The passage between the esophagus and the stomach is guarded by the cardiac sphincter or cardia. This exists in a permanently contracted state, or tonus, superimposed on which from time to time are rhythmic alternations of contraction and relaxation. This tonus is never very pronounced. In man it is said that a water pressure of from 2 to 7 cm. applied to the esophageal side of the sphincter will drive air or water into the stomach, this pressure being less than that of a column of fluid filling the thoracic esophagus in the erect position. During repeated deglutition the tonus becomes less and less marked, and after a number of swallows the sphincter may become completely relaxed. When this relaxation disappears, however, the sphincter becomes more contracted than usual and remains so for a longer time.

The tonic condition of the sphincter is controlled by the vagus nerve, stimulation of which causes relaxation with an after-effect of strong contraction. Mechanical or chemical stimulation of the lower end of the esophagus increases the tonus of the sphincter. Forcing of the sphincter from the stomach side requires a higher pressure than from the esophageal. Eructation of gas, for example, does not take place until intragastric pressure has risen to about 25 cm. of water. In deep anesthesia, however, intragastric pressure may rise considerably higher without forcing the sphincter.

In animals fed with starch paste impregnated with subnitrate of bismuth and then examined by means of the x-rays, the variation in degree

of tone of the sphincter has been observed to be responsible for *occasional regurgitation of some of the gastric contents* into the esophagus up to the level of the heart or even to the base of the neck. The presence of the gastric contents in the esophagus starts a peristaltic wave, which pushes the material back again into the stomach. This peristaltic wave starts in the absence of any other phases of the deglutition process, indicating that it has been excited by the presence of the material in the esophagus itself, and belongs, therefore, to the lower order of peristaltic wave, as seen in the intestines but not in the upper half of the esophagus. Regurgitation of food into the esophagus occurs only when the intragastric pressure is fairly high. It may last for a period of from twenty to thirty minutes after the meal is taken, and disappears when the tonus of the sphincter becomes increased as a result of the presence in the gastric contents of free hydrochloric acid.

Much information has been secured by listening with a stethoscope to *the sounds caused by swallowing* and by observing with the x-ray the *shadows* produced along the course of the esophagus when food impregnated with bismuth subnitrate is taken. When a solid bolus is swallowed only one sound is usually heard, but with liquid food there are two, one at the upper end, due to the rush of the fluid and air, and the other at the lower end (heard over the epigastrium), four or six seconds later, due to the arrival here of the peristaltic wave with the accompanying opening of the cardiac sphincter and the escape of the fluid and air into the stomach. Sometimes, when the person is in the horizontal position, this second sound may be broken up into several, indicating that, unassisted by gravity, the fluid does not so readily pass through the sphincter. The x-ray shadows yield results in conformity with the above. After swallowing milk and bismuth, for example, the shadow falls quickly to the lower end of the esophagus and then passes slowly into the stomach. When the passage of a solid bolus is watched by the x-ray method, its rate of descent will be found to depend on whether or not it is well lubricated with saliva; if not so, it may take as long as fifteen minutes to reach the stomach; if moist, but from eight to eighteen seconds.

Vomiting

Vomiting is usually preceded by a feeling of sickness or nausea, and is initiated by a very active secretion of saliva. The saliva, mixed with air, accumulates to a considerable extent at the lower end of the esophagus, which it distends. A forced inspiration is now made, during the first stage of which the glottis is open so that the air enters the lungs, but later the glottis closes so that the inspired air is sucked into the

esophagus, which, already somewhat distended by saliva, now becomes markedly so. The abdominal muscles then contract so as to compress the stomach against the diaphragm and, simultaneously, the cardiac sphincter relaxes, the head is held forward and the contents of the stomach are ejected through the previously distended esophagus. The compression of the stomach by the contracting abdominal muscles is assisted by an actual contraction of the stomach itself, as has been clearly demonstrated by the x-ray method. After the contents of the stomach itself have been evacuated, the pyloric sphincter may also relax and permit the contents (bile, etc.) of the duodenum to be vomited.

The act of vomiting is controlled by a center located in the medulla, and the *afferent fibers* to this center may come from many different regions of the body.. Perhaps the most potent of them come from the sensory nerve endings of the fauces and pharynx. This explains the tendency to vomit when the mucosa of this region is mechanically stimulated. Other afferent impulses come from the mucosa of the stomach itself, and these are stimulated by *emetics,* important among which are strong salt solution, mustard water and zinc sulphate. Certain other emetics, particularly tartar emetic and apomorphine, act on the vomiting center itself, and can therefore operate when given subcutaneously. Afferent vomiting impulses also arise from the abdominal viscera, thus explaining the vomiting which occurs in strangulated hernia, and in other irritative lesions involving this region. X-ray observations have been made on the movements of the stomach of cats after the administration of apomorphine (Cannon). The first change observed is an inhibition of the cardiac end of the stomach, which becomes a perfectly flaccid bag. About the midregion of the organ, deeper contractions then start up, which sweep from the pylorus, each contraction stopping as a deep ring at the beginning of the vestibule, while a slighter wave continues. A very strong contraction at the incisura angularis finally develops and completely divides the gastric cavity into two parts. On the left of this constriction the stomach remains completely relaxed, but at the right of it waves continue running over the vestibule. It is while the stomach is in this condition that the sudden contraction of the diaphragm and abdominal muscles shoots the cardiac contents into the relaxed esophagus. As these jerky contractions are continued, the gastric walls seem to reacquire their tone.

CHAPTER LII

THE MECHANISMS OF DIGESTION (Cont'd)

THE MOVEMENTS OF THE STOMACH

The Character of the Movements

Even from the earliest days it has been recognized that the stomach performs two important functions: (1) receiving the swallowed food and then discharging it slowly into the intestine, and (2) initiating the chemical processes of digestion. In order to understand the mechanism by which the stomach collects and then discharges the food, it is necessary first of all to recall certain anatomic facts concerning the organ, and for this purpose it is most convenient to accept the description given by Cannon, which is illustrated in the accompanying figure. The organ is divided into a cardiac and a pyloric portion by a deep notch in the lesser curvature, called the incisura angularis. The cardiac portion is further subdivided into two by the cardiac orifice. The part which lies, in man, above a line drawn horizontally through the cardia is the fundus. The part lying between the fundus and the incisura angularis is known as the body of the stomach, which, when full, has a tapering shape. The pyloric portion lying on the right of the incisura angularis is further divided into two parts: the pyloric vestibule and the pyloric canal, the latter of which lies next the pyloric sphincter and in man measures about 3 cm. in length (see Fig. 154).

The filled stomach of a person standing erect is so disposed that the greatest curvature forms its lowest point, which may be considerably below the umbilicus. As digestion proceeds and the stomach empties, the greater curvature becomes gradually raised, so that ultimately the pylorus comes to be the most dependent part of the stomach. From these and many other observations it is certain that the emptying of the stomach does not at all depend on the operation of the force of gravity. Indeed, that this can not be the case is perfectly clear when we consider the disposition of the stomach in quadrupeds.

Exact observation on the movements which the stomach performs from the time it is filled with food till it empties, have been made by the x-ray method, first introduced by Cannon.[12] The method consists in feed-

ing the animal with food that has been impregnated with bismuth subnitrate, then exposing him to the x-ray and either taking instantaneous photographs of the shadows or observing them by means of a fluorescent screen. The descriptions of the original observations made by Cannon

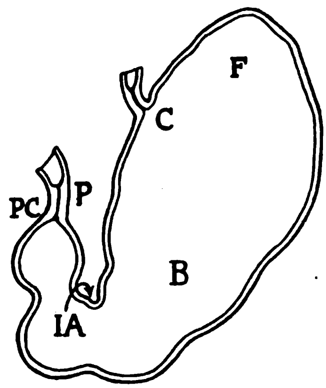


Fig. 154.—Schematic outline of the stomach. At *C* is the cardia; *F*, fundus; *IA*, incisura angularis; *B*, body; *PC*, pyloric canal; *P*, pylorus. (From Cannon.)

on the stomach of the cat have been so little modified by observations on man that we may take them as a convenient type. In the accompanying figure (Fig. 156) the outline of the shadow cast by the stomach is shown at intervals of an hour each during digestion. Soon after the

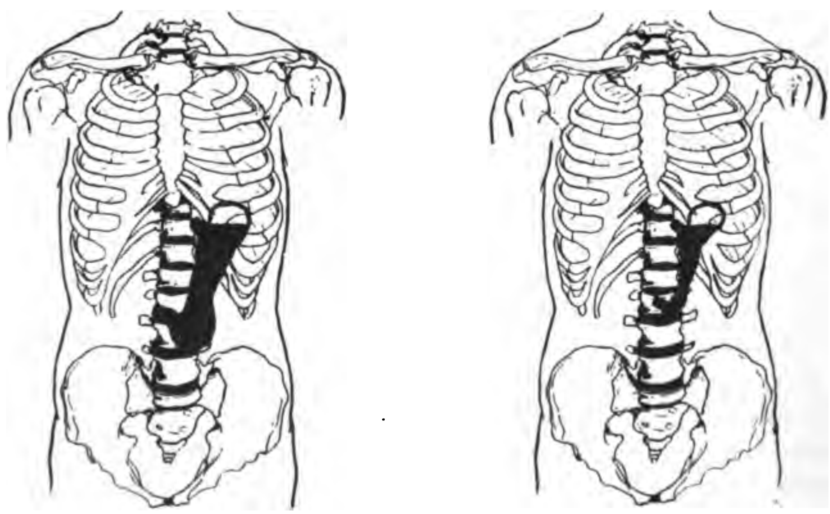


Fig. 155.—Diagrams of outline and position of stomach as indicated by skiagrams taken on man in the erect position at intervals after swallowing food impregnated with bismuth subnitrate. *A*, moderately full; *B*, practically empty. The clear space at the upper end of the stomach is due to gas, and it will be noticed that this "stomach bladder" lies close to the heart. (From T. Wingate Todd.)

stomach has become filled, peristaltic waves are seen to take their origin about the middle of the body of the stomach, and to course towards the pylorus. Above the region at which these waves originate—that is, the cardiac half of the body of the stomach and all of the fundus—there are no waves, but as digestion proceeds the walls slowly and steadily contract on the mass of food. This so-called cardiac pouch does not, however, diminish in size so rapidly as the part of the body of the stomach over which the peristaltic waves are passing. The circular fibers of the walls of this part of the stomach—sometimes called the gastric tube—contract tonically, so that it becomes tubular in form, with the full cardiac pouch at the left and above and the pyloric portion at the right. The latter portion meanwhile does not diminish much in size, although the peristaltic waves traveling over it are very pronounced. As will be clear from the figure, these changes in outline go on until the cardiac pouch has become practically empty and the food has been all moved along the now tubular portion of the body into the pyloric vestibule.

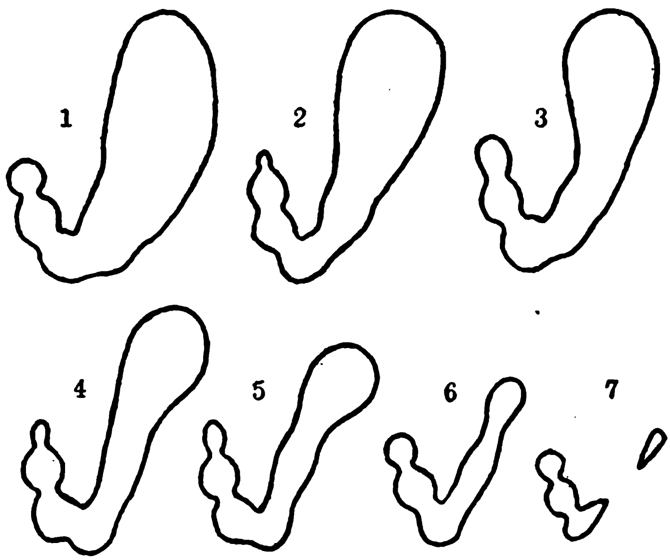


Fig. 156.—Outlines of the shadows cast by the stomach at intervals of an hour each after feeding a cat with food impregnated with bismuth subnitrate. (From Cannon.)

From this description it is evident that the function of the cardiac end is to serve as a reservoir for the food, which, by a slow contraction of the walls, is gradually delivered into the gastric tube, where by peristalsis it is carried towards the pyloric vestibule.

The time required for the peristaltic waves to travel from their place of origin to the pylorus is considerably longer than the interval between

the waves, so that several of these are always seen on the stomach at the same time. They sometimes become so pronounced in the pyloric region, especially in a half-empty stomach, that they appear almost to obliterate the cavity. They always stop at the pylorus, never going on to the duodenum. The rate of recurrence of the waves varies somewhat in different animals, being about six per minute in the cat and about three in man. Their initiation does not seem to depend on the presence of acid in the gastric contents, for, when food is introduced into the stomach, they do not wait for the gastric contents to become acid in reaction (see page 482). Nevertheless, acid does seem somewhat to stimulate the depth and frequency of the waves, and they recur oftener with carbohydrate than with fatty food.

The pressure in the stomach contents—*the intragastric pressure*—is low and constant at the cardiac end and fairly high and variable in the pyloric end (in the former from 6 to 8 cm. of water, and in the latter from 20 to 30). Constancy of pressure in the cardiac end indicates that the stomach wall must adapt itself very promptly to the amount of food in the organ. The higher and more variable pressure in the pyloric end is, of course, due to the peristaltic waves, and it is interesting to note that it is sufficient to propel the gastric contents through the pylorus for several centimeters into the duodenum.

The Effect of the Stomach Movements on the Food

This has been studied: (1) by dividing the food into portions that are differently colored and, after some time, killing the animal, freezing the stomach and making sections of it (see Fig. 157); (2) by making little pellets of bismuth subnitrate with starch and observing their behavior under the x-rays; or (3) by removing samples of the stomach contents by means of a stomach tube (Rehfuss tube) inserted so that its free end lies in either the cardiac or the pyloric region. By the first of the above methods it has been found that the first mouthfuls of food lie along the greater curvature, where they form a layer over which that subsequently swallowed accumulates, with the last portions next the cardia. The pepsin and hydrochloric acid of the cardiac end, therefore, act soonest on the first swallowed portion of a meal, and the more recently swallowed central masses are not affected by the secretions for some time, so that opportunity is given for the saliva mixed with the food to develop its digestive action.

As has been shown by removing the stomach contents with a tube at various periods after feeding with starchy food, considerable amylolysis may occur for some time. When separate samples are removed in this way from the cardiac and pyloric parts, it has been found that after

half an hour the contents of both have about the same percentage of sugar, but that for some time after this interval the cardiac contents contain considerably more sugar than the pyloric. Later the percentages of sugar again become about equal, no doubt on account of diffusion. The diastatic action in the fundus is finally brought to an end when the contents become completely permeated by the hydrochloric acid. In this connection it is worthy of note that the addition of hydrochloric acid up to the point of neutrality greatly accelerates the rate of diastatic digestion.

As the outer layers of food in the stomach become partly digested on account of the action of the pepsin and hydrochloric acid, the food is slowly pressed into the active right half of the stomach, where by the action of the peristaltic waves it is moved on to the pyloric vestibule. By observing the x-ray shadows cast by two pellets of bismuth subnitrate it has been noted by Cannon that, as the peristaltic wave approaches

Fig. 157.—Section of the frozen stomach (rat) some time after feeding with food given in three differently colored portions. (From Howell's *Physiology*.)

a pellet, it causes it to move forward more rapidly for a short distance, but soon overtakes it and in doing so causes the pellet to move back a little towards the fundus. This backward movement is less than the forward movement, so that after the wave has passed, the position of the pellet is a little forward of that which it would have occupied had there been no wave. The behavior of the pellet, and, therefore, of the stomach contents, is very like that of a cork floating at the edge of the sea; as each wave approaches, it hurries the cork on a little, but after its passage the cork recedes again until the second wave carries it still a little farther forward. As the peristaltic wave approaches the pyloric vestibule and becomes more powerful its effect on the pellets becomes more marked. They are carried rapidly along this part of the stomach, until the pylorus is reached. If this remains closed, they are shot back into the vestibule. From nine to twelve minutes may elapse before they are transferred to the pylorus from the place where they are first affected by the peristaltic wave.

These observations made on cats and other laboratory animals no doubt also apply in the case of man. Removal of the contents of the cardiac and pyloric regions separately with a stomach tube after feeding with a test meal part of which was colored with carmine or charcoal, has shown that none of the coloring material was present in the contents of the pyloric end up to twenty minutes or so after the food had been taken. It then appeared but at first only in traces. Another important distinction between the food in the two portions of the stomach relates to its consistency. In the pyloric end it is semifluid and homogeneous in character; in the cardiac end, on the other hand, it is a lumpy, rather incoherent mass.

The gastric movements must greatly *facilitate the digestive processes* in the stomach. In the cardiac part the undisturbed condition of the food will, as we have seen, facilitate the digestive action of ptyalin, whereas in the body of the stomach the peristaltic waves, besides moving the food onward, will tend to bring fresh portions of mucous membrane and food in contact, so that the latter becomes more thoroughly mixed with the pepsin and hydrochloric acid. In the pyloric part, where no hydrochloric acid is secreted, the contents, already sufficiently acid in reaction, become more thoroughly churned up with the local pepsin secretion, so that proteolytic action progresses very rapidly.

The peristaltic waves also *facilitate absorption from the stomach* of such substances as glucose in concentrated solution and, probably, of hydrolyzed protein; water, however, is not absorbed. One effect of such absorption is the production of gastrin, which we have seen is the hormone concerned in maintaining the gastric secretion after the psychic flow. The fact that the mucosa of the vestibule has, relatively to the cardiac end, few secreting glands is in harmony with the view that absorption is an important function of this part of the stomach.

THE EMPTYING OF THE STOMACH

The Control of the Pyloric Sphincter

When digestion has proceeded far enough in the stomach to bring the food into a homogeneous, souplike fluid (chyme), portions of this, as they are driven against the pyloric sphincter by the peristaltic waves, instead of being returned as an axial stream into the stomach, are ejected into the duodenum.

We must now consider the mechanism by which the pyloric sphincter opens to permit the passage of the chyme. Bombardment by the peristaltic waves is evidently not the cause of its opening, for, as we have

seen, many such waves may arrive at it without this result. Since it is evidently in order that the intestine may not suddenly become overwhelmed with large masses of food that the pylorus only occasionally opens, it might be thought that its opening depends upon the distention of the upper part of the intestine. It is true that excessive distention of the upper part of the intestine does hold the pyloric sphincter closed, but this can not be the physiologic stimulus, because considerable quantities of chyme are never found here.

The first clue to the real nature of the mechanism was afforded by observing the behavior of the sphincter when solutions are introduced into the duodenum through a fistula. Acid solutions were found to cause a complete inhibition of gastric evacuation, whereas alkaline solutions had no effect. This difference indicates that acids in contact with the duodenal mucous membrane reflexly excite contraction of the sphincter, and that it relaxes only after the acid has become neutralized by mixing with the pancreatic juice and bile.

On account of the great importance of the pyloric mechanism in insuring that the chyme shall enter the intestine only in such quantities that it can be properly acted upon by the intestinal digesting juices, it will be of interest to consider briefly some of the experimental observations by which this mechanism has been studied. We may consider first the evidence that *acid on the stomach side of the pylorus causes a relaxation of the sphincter:* (1) When carbohydrate food is fed, it ordinarily leaves the stomach fairly rapidly, but if its acid-absorbing power is increased by mixing it with sodium bicarbonate, exit from the stomach is greatly delayed. (2) Proteins ordinarily leave the stomach more slowly than carbohydrates, but if acid proteins are fed, their exit is much more rapid. (3) If a fistula is made into the pyloric vestibule through which some of the contents can be removed, it will be found that just prior to the opening of the pyloric sphincter, a distinctly acid reaction develops in the food; and furthermore if acid solutions are injected through this fistula, they cause the pyloric sphincter to open, whereas alkalies retard its opening. (4) A similar effect of acid in opening the sphincter can be demonstrated by applying it to the pyloric mucosa of an excised stomach kept alive in oxygenated Ringer's solution.

The evidence that *acid on the duodenal side causes closure of the sphincter* is as follows: (1) When acid is placed in the duodenum through a fistula, the sphincter will not open; (2) when the pancreatic and bile ducts are ligated, the stomach empties much more slowly than normally; and (3) the discharge of protein is considerably hastened if the pylorus is sutured to the intestine below the duodenum. After such an operation it was observed that the protein began to leave the stomach through

the pyloric sphincter about the same time as normally, but the subsequent evacuation was very much accelerated, because no acid came in contact with the duodenal mucosa. Water and egg white may leave the stomach independently of any acid reflex control of the pylorus. By observations made through a duodenal fistula it has been found that, after a quantity of water has been swallowed, most if not all of it very soon enters the duodenum in a more or less continuous stream. It is no doubt on this account that drinking contaminated water is especially dangerous on an empty stomach.

The *nervous pathway* through which these acid reflexes take place has been shown to be the myenteric plexus. Indeed, the whole mechanism is quite analogous with that which we shall see occurs in the intestine during peristalsis: the stimulus, that is, the acid, causes a contraction of the gastric tube behind it and a dilatation in front.

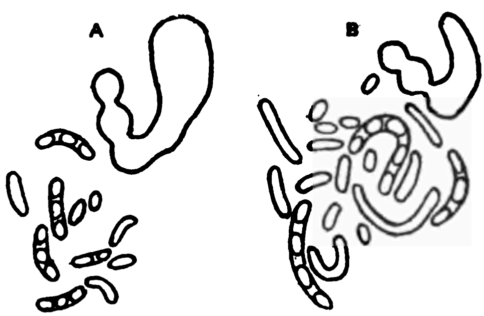


Fig. 158.—Outlines of shadows in abdomen obtained by exposure to x-rays 2 hours after feeding with food containing bismuth subnitrate. The food in *A* was lean beef, and in *B* boiled rice. The smaller size of the stomach shadow and the much greater total area of the intestinal shadows in *B* than in *A* show that carbohydrate leaves the stomach earlier than protein. (From Cannon.)

Rate of Emptying of Stomach

The relationship of these facts *to the rate at which different foodstuffs leave the stomach* is very readily explained. The method for investigating this problem, which again we owe to Cannon, consists in feeding animals with a strictly uniform amount of different foods made up, as nearly as possible, of equal consistency and containing bismuth subnitrate in the proportion of 5 gm. to each 25 c.c. By feeding such mixtures to cats previously starved for twenty-four hours, and examining the abdomen by the x-ray at regular intervals, the shadows cast by the food after passage into the intestine can be outlined on tracing paper, and the total length* measured (Fig. 158). In taking this as an estimate of the amount of food in the intestine, several errors are no doubt incurred

*This is permissible since the shadows are practically all of the same width.

on account of the crossing and foreshortening of the loops, etc., but, as their constancy testifies, there is no doubt that the results are sufficiently close for the purpose of finding out how quickly food gains access to the small intestine; and the method has a great advantage over all others in that digestion is allowed to proceed practically without interruption. The points we have to determine are: (1) when the food first leaves the stomach; (2) the rate at which different foods are discharged; (3) the time required for the passage through the small intestine.

Let us consider first of all the results obtained *by feeding with practically pure fat or carbohydrate or protein.* By plotting the length of the shadows in centimeters along the ordinates, with hours along the abscissæ, curves such as those shown in Fig. 159 have been secured. When fats were fed (dash line in chart), the discharge began rather slowly, and continued at a slow rate. Even after seven hours some fat still remained in the stomach, and at no time was any large quantity

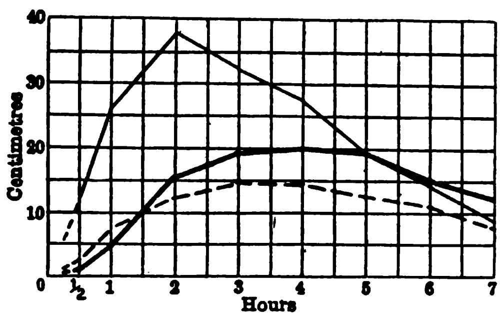


Fig. 159.—Curves to show the average aggregate length of the food masses in the small intestine at the designated intervals after feeding. The curve for various fat foods is in the dash line, for protein foods in the heavy line, and for carbohydrate foods in the light line. (From Cannon.)

present in the intestine, indicating that almost as quickly as it is discharged into this part of the gastrointestinal tract fat becomes digested and absorbed. The discharge of carbohydrates was quite different (light line in chart); it began often in ten minutes, and soon became abundant, reaching a maximum, as a rule, at the end of two hours, after which it fell off, the stomach being empty in about three hours. Protein left at a rate intermediate between that for fats and that for carbohydrates (heavy line). Little left before the first half hour; the curve then slowly rose, attaining a maximum in about four hours, and then gradually declining at about the same rate as it rose. It is interesting to note that at the end of half an hour about eight times as much carbohydrate had left the stomach as protein; at the end of an hour, five times as much.

These results are clearly dependent upon the rates at which the different foodstuffs assume an acid reaction in the stomach. Carbohydrate

has no combining power for acids, so that the acid secreted with the psychic juice remains uncombined and on gaining the pyloric vestibule excites the opening of the sphincter. Protein, on the other hand, as is well known, absorbs considerable quantities of free hydrochloric acid, so that for some considerable time after it is taken, none of the acid exists in a free state. Fats owe their slow discharge partly to inhibition of gastric secretion, and partly to the longer time it takes for them to become neutralized in the duodenum, because of the fatty acid split off by the action of lipase.

Interesting observations have also been made on the rate of discharge when various *combinations of foodstuffs* were fed. This has been done by feeding one foodstuff before the other, or by mixing the foodstuffs. When carbohydrates were fed first and then protein, the discharge began much earlier than with protein alone, because the carbohydrate food first reached the pyloric vestibule (see page 454). However, at the end of two hours, when the carbohydrate curve should begin to come down, it remained high, indicating that the protein had by this time reached the pylorus and was being discharged at its own rate. When the meat was fed before the carbohydrate, the curve to start with was exactly like that for protein, becoming, however, considerably heightened later when the carbohydrate reached the pyloric vestibule. The presence of protein near the pylorus, therefore, distinctly retards the evacuation of carbohydrate from the stomach. These facts, it will be remarked, all fit in admirably with the observations which we have already detailed concerning the disposition of food in the stomach.

When mixtures of equal parts of different foods were fed, the results indicated that the emptying of the stomach occurred at a rate which was intermediate between those of the foods taken separately. Mixing protein with carbohydrate, for example, accelerated the rate at which protein left, and mixing fats with protein caused the protein to leave the stomach considerably more slowly than if protein alone had been fed.

Influence of Pathologic Conditions on the Emptying

An important surgical application of these facts concerns *the behavior of food after gastroenterostomy.* It has been thought that this operation would cause the food to be drained from the stomach into the intestine and thus leave the region of the stomach between the fistula and the pylorus inactive. This assumption is based on the idea, which we have seen to be erroneous, that gravity assists in the emptying of the stomach. As a matter of fact, it has been found that, if the gastroenterostomy is made when there is no obstruction at the pylorus, the chyme takes its

normal passage through the sphincter and, almost without exception, none leaves by the fistula. When the pylorus is partly occluded, the food sometimes passes in the usual way, and sometimes by the stomach. The cause for this predilection for the pyloric pathway depends on the pressure conditions in the gastric contents. Gastroenterostomy, therefore, is efficient only when gross mechanical obstruction exists at the pylorus. The operation should never be performed in the absence of demonstrable organic pyloric disease.

Another objection to gastroenterostomy in the presence of a patulous pyloric sphincter rests on the fact that the food, after passing the sphincter and moving along the intestine, may again enter the stomach through the fistula. This is most likely to occur when the stomach is full of food, for under these conditions the stretching of its walls separates the edges of the opening, the intestine being drawn taut between the edges, so that the opening between the stomach and the intestine assumes the form of two narrow slits, which act like valves permitting the food to enter but preventing its escape from the stomach. Only seldom under these circumstances can any food pass into the intestine beyond the stomach opening. Repeated vomiting after gastroenterostomy has been observed in experimental animals only when obstructive kinks or other demonstrable obstacles were present in the gut, the obstruction being located in that part of the intestine beyond its attachment to the stomach.

When the pyloric obstruction is complete, food must, of course, leave by the fistula, digestion by the pancreatic juice and bile being still carried on because of the fact that for a considerable distance down the intestine, secretin, which we have seen is essential for the secretion of these fluids, is still produced by the contact of the acid chyme with the intestinal mucosa. Further provision for adequate digestion of food in such cases is secured, as some of the food after leaving the fistula passes back for some distance into the duodenum, where, however, it soon excites peristaltic waves, which again carry it forward. This insures thorough mixing with the digestive juices. From their experimental experience Cannon and Blake[13] recommend that, when the fistula has to be made, it should be as large as possible and near the pylorus, and that the stomach afterwards should not be allowed to become filled with food. To avoid kinking of the gut, they also recommend that several centimeters of the intestine should be attached to the stomach distal to the anastomosis.

The effect of *hyperacidity* of the contents on the emptying of the stomach has been studied by feeding animals with potatoes containing varying percentages of hydrochloric acid. With an acidity of 0.25 per

cent, the rate of discharge was increased, but it became slower when the acidity rose to 1 per cent. With an acidity of 0.5 per cent, the rate of discharge was about the normal. Hyperacidity, therefore, causes a retardation of the emptying of the stomach.

The *consistency* of the food appears to have little influence on its rate of discharge from the stomach—at least in the case of potatoes. Dilution of protein food, however, increases the rate. Distinctly hard particles in the food retard the stomach evacuation.

There is usually a considerable amount of *gas* in the part of the stomach above the entrance of the cardia, on account of which this part of the stomach has sometimes been called the stomach bladder. In the upright position this gas forms a bright area in the x-ray plate (Fig. 155), but when the person reclines it spreads to a new location. Its presence may influence gastric digestion by preventing the contact of the food with the mucous membrane, and by interfering with the efficiency of the peristaltic waves in moving the food. Considerable gas therefore retards the emptying of the stomach, as has been shown experimentally by x-ray observations on animals fed with the standard amount of food followed by the introduction of air. It was noted that the air did not diminish the frequency or strength of the peristaltic waves, but that these could not efficiently act on the food. When along with gas there is also atony of the stomach walls, the retardation in the discharge will, of course, be still more pronounced. The temperature of the swallowed food does not appear to have much influence on the stomach movements or on the the rate of discharge from the organ.

CHAPTER LIII

THE MECHANISMS OF DIGESTION (Cont'd)

THE MOVEMENTS OF THE INTESTINES

The length of the small intestine and the size of the cecum of the large intestine vary considerably in different animals. In the carnivora, such as the cat, the small intestine is relatively short; in the herbivora, relatively long. Thus, it is three times the length of the body in the cat, and four to six times in the dog; whereas in the goat and sheep, it may be nearly thirty times the length of the body. In the carnivora the cecum is either absent or rudimentary, whereas in those herbivora which do not have a divided stomach the cecum is very large and sacculated, as is also the colon. The reason for the great size in herbivora is that practically the whole of the digestion of cellulose takes place in this part of the gut. This digestion, as we shall see later, does not depend on any secretion poured forth by the animal itself, but upon the action of bacteria and of certain enzymes (cytases) that are taken with the vegetable food.

Movement of the Small Intestine

The movements of the small intestine have been studied (1) by the bismuth subnitrate and x-ray method, (2) by observing them after opening the abdomen of an animal submerged in a bath of physiologic saline at body temperature, (3) by observing the changes in pressure produced in a thin-walled rubber balloon inserted in the lumen of the gut and connected with a recording tambour (Fig. 160), and (4) by excising portions of the intestine and keeping them alive in a bath of saline solution at body temperature, through which oxygen is made to pass.

The Segmenting Movements

When a suitably fed animal is placed on the holder for examination by the x-ray method, no movement in the intestinal shadows is generally observed for some time. The first movement to appear is the breaking of one of the columns of food into small segments of nearly equal size. Each of these segments again quickly divides, and the neighboring halves suddenly unite to form new segments, and so on, in a manner

which will be made clear by consulting Fig. 161. This ***rhythmic segmentation,*** as Cannon has called it, continues without cessation for more than half an hour, and the food shadow does not meanwhile seem to change its position in the abdomen to any extent. The splitting up of the segment and the rushing together of the neighboring halves proceed as a rule with great rapidity; thus, if we count the number of different seg-

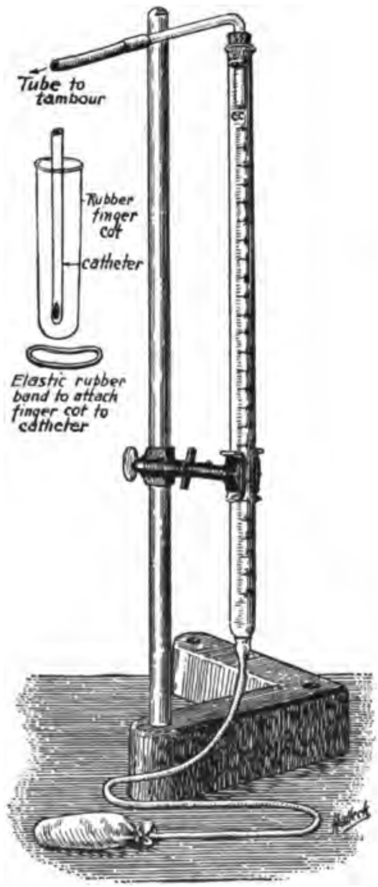


Fig. 160.—Apparatus for recording contractions of the intestine. (From Jackson.)

ments during a definite period, we may find the rate of division in the cat to be as high as 28 or 30 a minute. In man the divisions occur at a frequency of approximately 10 per minute, which corresponds to the frequency with which sounds can be heard when the abdomen is auscultated. Although half an hour is the period which this process usually occupies, it may last considerably longer. In certain animals, such as the rabbit, segmenting movements have not been observed, but instead

of them a rhythmic to-and-fro shifting of the masses of food along the lumen of the gut, rapidly repeated for many minutes.

When the intestines are floated out in a warm bath of saline solution, it is seen that the rhythmic segmentation is caused by narrow rings of contraction. Under such conditions also it is often noted that the loops of intestine sway from side to side. The balloon method also reveals the presence of slight waves of contraction that pass rapidly along the gut, and follow each other at the rate of twelve to thirteen per minute. Both of the muscular coats of the intestine are involved, and it is believed that the contractions are responsible not only for the pendular movements but for the rhythmic segmentation observed by the x-ray method. According to this view these movements are constantly passing along the intestine, and become exaggerated by the mechanical stimulus which is offered by the masses of food to such an extent that they divide the masses into portions. The evidence for this belief rests on the fact that

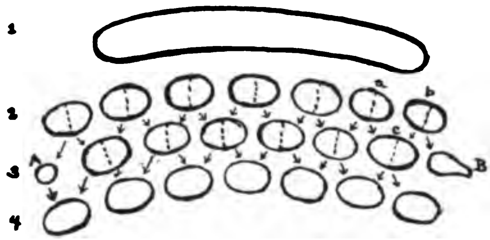


Fig. 161.—Diagrammatic representation of the process of segmentation in the intestine. An unbroken shadow is shown in *1* and its segmentation in *2*. The dotted lines across each mass show the position of division and in *3* is shown how new masses are formed by the split portions coming together. (From Cannon.)

when the contraction is studied by the balloon method, it becomes marked over the middle of the balloon, where the greatest tension exists.

Several functions can be assigned to these movements. They cause intimate mixture of the food with the digestive juices, and by bringing ever new portions of food in contact with the mucosa, they encourage absorption. They also have an important massaging influence on the blood and lymph in the vessels of the intestinal walls. Indeed, the passage of lymph from the lacteals into the mesenteric lymphatics seems to depend very largely upon these movements.

The Peristaltic Movements

The other movement observed in the small intestine is that known as the *peristaltic wave*. It occurs in two forms: (1) as a slowly advancing contraction (1 to 2 cm. per minute), preceded by an inhibition of the walls, and proceeding only through a short distance in a coil (4 to 5 cm.); and

(2) as a swift movement called the *peristaltic rush,* which sweeps without pause for much longer distances along the canal.

Further analysis of the peristaltic wave can readily be made by the balloon method (Fig. 162). If the gut is pinched above the balloon, a marked relaxation occurs over it, and this relaxation extends for about two feet down the intestine. If, on the other hand, the gut is pinched a little below the situation of the balloon, a long-continued contraction occurs over the latter. The conclusion that we may draw from this result is that the stimulation of the gut causes contraction above the point of the stimulus and relaxation below, this being known as "the law of the intestine"—(Bayliss and Starling). We have seen that it applies also in the case of the cardiac and pyloric sphincters.

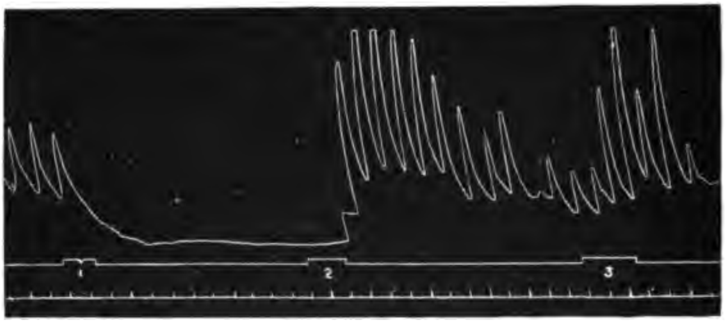


Fig. 162.—Intestinal contractions (balloon method) after excision of the abdominal ganglia and section of both vagi. Mechanical stimulation above (*1*) and below (*2*) the balloon causes relaxation and contraction respectively. (From Starling.)

The Physiologic Nature of the Rhythmic and Peristaltic Movements

Interesting information in this connection has been gained by observation of the behavior of the movements after the application of drugs to the gut or after cutting the nerve supply. The rhythmic movements are not affected by the application of nicotine or cocaine. Since these drugs paralyze nervous structures it has been concluded that the rhythmic movements are myogenic in origin. The question is not a settled one, however, for it has been found by Magnus that, although strips of the longitudinal muscle, isolated in oxygenated saline solution, will continue to beat, they do not do so if the adherent Auerbach's plexus of nerves is stripped off from them. The nature of the peristaltic contractions is more definite; they must clearly depend upon a local nervous structure, since they are paralyzed by the application to the gut of cocaine or nicotine. This local nervous system no doubt also resides in Auerbach's plexus, which must therefore be considered as complex enough to be (see

page 796) endowed with the power of directing nervous impulses so as to bring about relaxation of the gut in front of the stimulus and contraction over it.

Nervous Control of Movements

The influence of the central nervous system on the intestinal movements has been studied by the usual methods of cutting and stimulating the extrinsic nerve supply. Through the splanchnic nerves tonic *inhibitory* impulses are conveyed to the intestine (except the ileocolic sphincter), for after these nerves are severed the movements become more distinct. Indeed, in many animals after opening the abdomen no intestinal movement can be observed until these nerves have been cut. Stimulation of the peripheral end of the nerve also inhibits any movement which may meanwhile be in progress. The impulses through the vagus nerve are of an

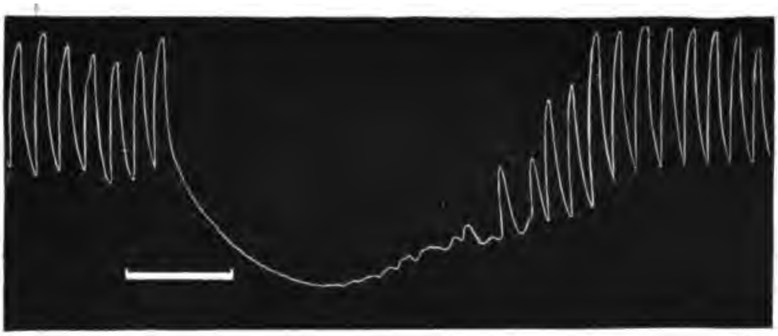


Fig. 163.—The effect of excitation of both splanchnic nerves on the intestinal contractions. (From Starling.)

opposite character. Section of these nerves has little effect, but stimulation causes contraction. (Figs. 163 and 164.)

By observing the rhythmic contractions of an isolated strip of the small intestine suspended in a bath of oxygenated saline solution at body temperature, it can readily be shown that the presence of even a minute trace of epinephrine is sufficient to produce complete inhibition of the movement. The parallelism between the effects of splanchnic stimulation and those of epinephrine injection is very significant, for in this way the marked inhibition of intestinal movement which occurs during fright may possibly be explained (see page 736).

The circular muscular coat of the last two or three centimenters of the ileum before it joins the cecum is definitely thicker than the rest of this coat, indicating that it has a sphincter-like action. This *ileocolic sphincter*, as it is called, opens when food is pressed against it from the ileum, but remains closed when food is pressed against it from the cecum.

It therefore obeys the law of the intestine. That it is physiologically distinct from the musculature of the rest of the ileum is indicated by the fact that the splanchnic and vagus nerves do not affect it in the same way; thus, stimulation of the splanchnic causes a strong contraction of the sphincter, whereas it is unaffected by stimulation of the vagus.

Peristalsis is much more rapid in the duodenum than in other parts of the small intestine. During the first stages of digestion, the food ordinarily lies mainly in the right half of the abdomen, and later in the left half. There is considerable variation in the time that elapses before it enters the colon. In the cat, carbohydrates reach this part of the gut in about four hours.

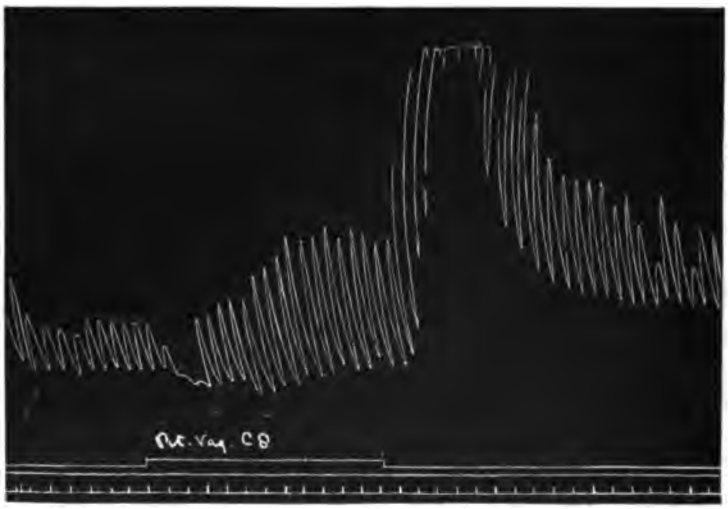


Fig. 164.—The effect of stimulation of right vagus nerve on the intestinal contractions. (From Starling.)

Movements of the Large Intestine

On account of the great differences which we have already seen to exist in the size and relative importance of the colon as a digestive organ in different classes of animals, it is not surprising that the movements observed are very different according to the dietetic habits of the animal. Apparently the movements are much the same in the cat as in man. As the food passes through the ileocolic sphincter into the cecum and accumulates there, it gradually sets up, by its pressure, a contraction of the muscular walls of the gut somewhere about the junction between the ascending and transverse colon. This wave of contraction then begins to travel slowly toward the cecum, without, however, being preceded by any relaxation of the wall of the gut, as is the case with a true

peristaltic wave. This first wave is soon followed by others, with the result that the food is forced up into the cecum, against the blind end of which it is crowded, being meanwhile prevented from passing into the ileum by the operation of the ileocolic sphincter and by the oblique manner in which the ileum opens into the cecum.

As the result of the distention of the cecum set up by these so-called antiperistaltic waves, a true coordinated peristaltic wave is occasionally initiated, and passes along the ascending colon preceded by the usual wave of inhibition. These waves, however, disappear before they reach the end of the colon, so that the food is again driven back by the so-

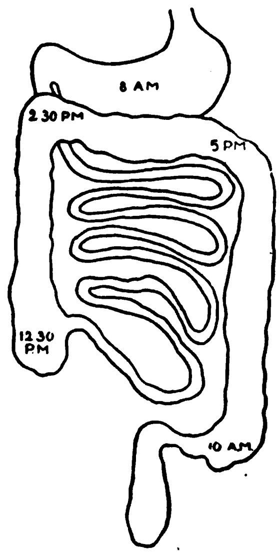


Fig. 165.—Diagram of time it takes for a capsule containing bismuth to reach the various parts of the large intestine.

called antiperistaltic waves. The effect of the movements is to knead and mix the intestinal contents, and thus encourage the absorption of water from them. The resulting more solid portions then collect toward the splenic flexure, and become separated from the remaining more fluid portion by transverse waves of constriction, which develop into peristaltic waves carrying the harder masses into the distal portions of the colon, where they collect chiefly in the sigmoid flexure. The descending colon itself is never distended with contents and merely serves as a tube for transferring the masses from the transverse colon to the sigmoid flexure. The time taken for a capsule of bismuth to reach the various parts of the large intestine is shown in Fig. 165.

After a certain mass has collected in the sigmoid flexure and rectum, the increasing distention causes a reflex evacuation of this portion of the

gut through centers located in the spinal cord. The impulses from these centers, besides contracting the rectum, etc., also coordinate the contraction of the abdominal muscles and the relaxation of the sphincter ani so as to bring about the act of defecation. By the skiagraphic method it has been found that the pelvic colon gradually becomes filled with feces from below upward, and that the rectum remains empty until just before defecation.

Effect of Clinical Conditions on the Movements

Observations of practical value have been made on the behavior of the peristaltic wave after various intestinal operations. *After an end-to-end anastomosis* of the gut, no evidence can be obtained by the x-ray method that any hesitation occurs in the movement of the shadows at the anastomosis. On the other hand, when a *lateral anastomosis* is established, stagnation of the food in the region of the junction may occur, this having been found, on opening the gut, to be caused by the accumulation of hair and undigested detritus at the opening between the opposed loops. Another objection to lateral anastomosis is the fact that in performing the operation a considerable amount of the circular muscle is cut, which interferes with peristaltic activity. Moreover, the end of the proximal loop beyond the opening is in danger of becoming filled up with hardened material, and the end of the distal loop may become invaginated and induce obstruction in the region of the anastomosis.

Observations have also been made by the x-ray method on the behavior of the intestinal contents following *intestinal obstruction.* It has been observed that, as the material collects in the gut just above the obstruction, strong peristaltic waves are set up, which move the food toward the obstruction so powerfully as to cause the walls of the canal in front to become bulged, until at last the pressure causes the contents to be squirted back through the advancing ring of peristaltic contraction. These waves were observed to succeed one another rapidly. When a portion of gut is reversed in position, the peristaltic waves continue to travel in their old direction toward the duodenum. The effect of this is to produce a partial obstruction at the upper end of the receptive gut.

The type of peristalsis known as the *peristaltic rush* can be induced experimentally in animals by intravenous injection of ergot. It probably also occurs in conditions of abnormal irritation of the gut in man, and is believed to be the characteristic activity of the gut after a strong purge.

CHAPTER LIV

HUNGER AND APPETITE

Hunger and appetite are distinct sensations, the former being definitely correlated with contraction of the empty stomach, and the latter being a complex of sensations operating in the nervous system along with memory impressions of the sight, taste, and smell of palatable food. Appetite is therefore a highly complex nervous integration, whereas hunger is a much simpler process. It is particularly with hunger that we shall concern ourselves at present.

When a thin-walled rubber balloon of proper size is placed in the stomach and connected by a rubber tube with a water, bromoform or chloroform manometer (made of wide glass tubing 1.5 cm. in diameter and provided with a suitable float on the free limb) a tracing may be taken of the movements of the stomach. For use on man the capacity of the balloon should be from 75 to 150 cubic centimeters. The record thus obtained when the balloon is placed in the empty stomach of a normal person shows four types of wave. Two of these may be discounted, being due to the arterial pulse and the respiratory movements. The third is known as the *tonus rhythm,* and is caused by tonic contractions of the fundus of the stomach of varying amplitudes and occurring at a rate which varies from 18 to 22 per second. The periods of tonus increase during the powerful rhythmic contraction to be immediately described. While these changes in tone are occurring, no subjective sensation of hunger is experienced. (See Fig. 167.)

The fourth and most significant type consists of powerful *rhythmic contractions,* alternating with periods of quiescence. These contractions occupy a period of about twenty seconds, and are superimposed upon the tonus rhythm. They gradually increase in amplitude and frequency; and, in the case of young and vigorous persons, may gradually pass into a condition of incomplete tetanus, after which they suddenly subside, leaving only a faint tonus rhythm. These rhythmic contractions are definitely associated with the sensation of hunger, and are more marked the more intense the sensation is. When tetanus occurs the hunger sensation is continuous, but it instantly disappears when the tetanus gives place to relaxation. When the contractions are comparatively feeble, the length of the period during which they occur is

about twelve minutes, being shorter than when the contractions are powerful, in which case the periods are always initiated by weaker contractions with long intervening pauses. Finally, the pauses disappear and the contractions become more and more pronounced until, as above mentioned, a virtual tetanus, lasting from two to five minutes, may supervene. The duration of the hunger period varies from one-half to one and a half hours, with an average of from thirty to forty-five minutes, and the number of individual contractions in a period varies from twenty to seventy. Between the hunger periods, intervals of from one-half to two and one-half hours of quiescence may supervene. (See Fig. 168.)

Similar contractions, often passing into incomplete tetanus, have been observed in the stomach of healthy *infants*, some of the observations having been made before the first nursing. The intervals of motor quiescence between the hunger periods are shorter than in adults. In observations made during sleep, it was observed that, when the contractions were very vigorous, the infant would show signs of restlessness and might awake and cry. As in the adult, the contractions are evidently associated with subjective sensations of hunger. Contractions of the empty stomach have also been recorded on a large variety of animals, including the dog, rabbit, cat, guinea pig, bird, frog and turtle. They vary somewhat in type in different animals.

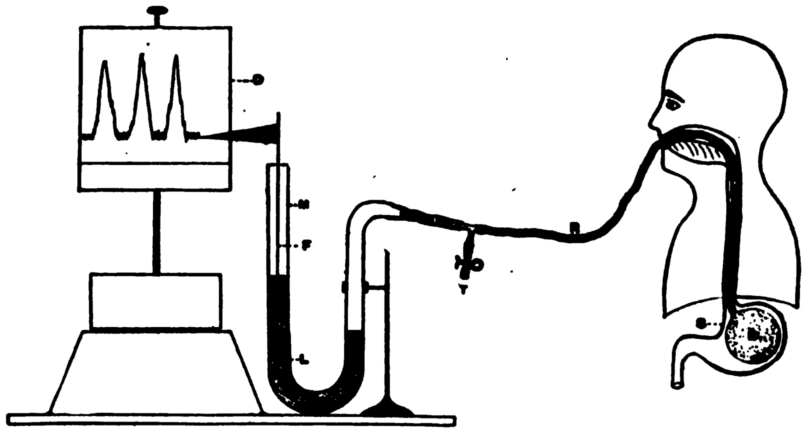


Fig. 166.—Diagram of method for recording stomach movements. *B*, rubber balloon in stomach. *D*, kymograph. *F*, cork float with recording flag. *M*, manometer. *L*, manometer fluid (bromoform, chloroform, or water). *R*, rubber tube connecting balloon with manometer. *S*, stomach. *T*, side tube for inflation of stomach balloon. (From Carlson.)

With regard to the time of onset of the tonus and hunger contractions, it has been observed that the only period during which the fundus is free of them is immediately after a large meal. After a moderate meal the tonus rhythm begins to appear in about thirty minutes. It gradually

increases in intensity, until by the time the stomach has nearly emptied itself the tonus has become conspicuous, and the stronger hunger contractions usually begin to appear. Superimposed upon those of the tonus rhythm, hunger pangs may appear in man *when the stomach still contains traces of food.*

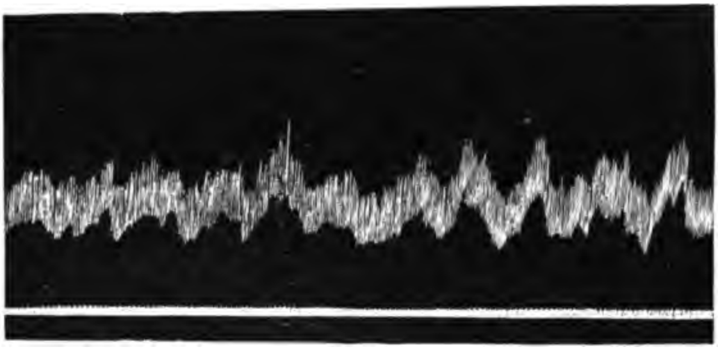


Fig. 167.—Tracing of the tonus rhythm of the stomach (man) three hours after a meal. (From Carlson.)

By studying the shadow of the outline of the stomach produced by having a person or animal swallow two balloons, one inside the other and with a paste of bismuth subnitrate between them, it has been observed that the weaker type of hunger contraction begins as a con-

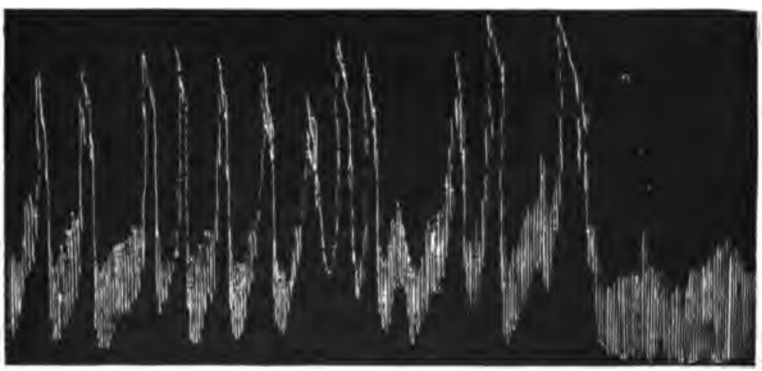


Fig. 168.—Tracings from the stomach during the culmination of a period of vigorous gastric hunger contractions. One-half original size. (From Carlson.)

striction involving the cardiac end of the stomach, and moving toward the pyloric end as a rapid peristaltic wave. When the contractions are very vigorous, this wave spreads so rapidly over the stomach that it is difficult to determine whether it really occurs as a very rapid peristalsis or as a contraction involving the fundus as a whole. These contractions

resemble very closely the movements that have sometimes been observed after a bismuth meal, and which have been thought by clinical observers to indicate a hyperperistalsis of the stomach. The fundus is therefore not entirely passive during digestion; for, although early in this act there may be no evidence of contraction, yet the contractions of the tonus rhythm may appear and become pronounced before the stomach is entirely empty. In other words, the digestion contractions of the filled stomach (see page 451) pass gradually over into the hunger contractions of the empty organ.

It appears that the stomach contractions produce the hunger sensations by causing stimulation of afferent nerve endings in the muscle layers of the viscus. Mere pressure on the mucosa itself does not originate such a sensation; thus, sudden distention of the balloon or rubbing the mucosa with the closed end of a test tube, inserted through a gastric fistula, did not cause any sensation of hunger, unless the stimulus was so strong as to cause a contraction of the musculature of the stomach.

It has been thought by some observers that, during hunger, contractions similar to those of the stomach also occur in the lower end of the esophagus. It is believed by Carlson, however, that these contractions are not at all responsible for the hunger sensation, although they may give rise to a feeling that something has stuck in the esophagus. Contractions of the intestine have also been observed in hunger, but it is doubtful whether they have anything to do with the cause of the hunger sensation.

REMOTE EFFECTS OF HUNGER CONTRACTIONS

It is well known that during hunger certain general subjective symptoms are likely to be experienced, such as a feeling of weakness and a sense of emptiness, with a tendency to headache and sometimes even nausea in persons who are prone to headache as a result of toxemic conditions. Headache is likely to be more pronounced or perhaps only present in the morning before there is any food in the stomach. These symptoms indicate that hunger contractions are associated with hyperexcitability of the central nervous system, and it is of considerable interest that objective signs of this association can be elicited. If the knee-jerk be recorded along with a record of the gastric contractions, it will be found that it is markedly exaggerated simultaneously with the strong hunger contractions of the empty stomach, this augmentation being greatest at the height of the stomach contractions, when the hunger pangs are most intense, and falling off again to normal when these disappear (Fig. 169). Further changes occurring during the hunger

period include an increase in the pulse rate and vasodilatation. By comparing plethysmographic tracings of the arm volume (see page 230) and stomach contractions, it has been found that the increase in volume occurs *pari passu* with the increasing tonus of the stomach, but that it begins to shrink before the stomach contraction has reached its maximum. Occasionally, however, as in acute hunger, a somewhat different relationship obtains, vasoconstriction being more prominent. During each hunger contraction there is also increased salivation, the degree of which varies with different individuals. This salivation is independent of the more copious "watering of the mouth" that accompanies the thought or sight of appetizing food.

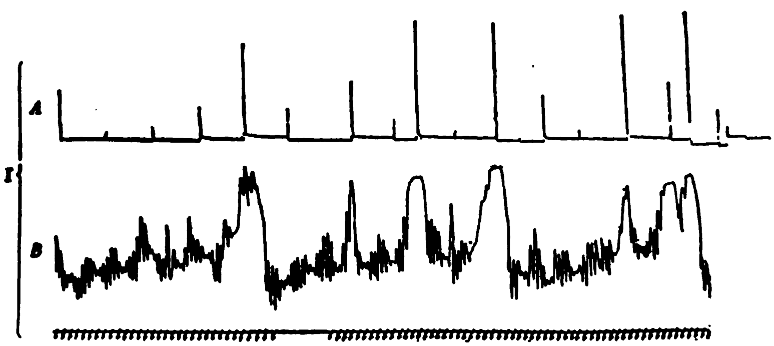


Fig. 169.—Showing augmentation of the knee-jerk (upper tracing) during the marked hunger contractions (lower tracing). (From Carlson.)

HUNGER DURING STARVATION

During enforced starvation for long periods of time, it is known that healthy individuals at first experience intense sensations of hunger and appetite, which last however only for a few days, then become less pronounced and finally almost disappear. It is of interest to know the relationship between these sensations and the hunger contractions in the stomach. This has been investigated by Carlson and Luckhardt, who voluntarily subjected themselves to complete starvation, except for the taking of water, for four days. During a great part of this time records of the stomach contractions were taken by the balloon method, and it was found that the tonus of the stomach and also the frequency and intensity of the hunger contractions became progressively more pronounced as starvation proceeded. Towards the end of the period it was also noted that incomplete hunger tetanus made its appearance where ordinarily, as in Carlson's case, this type of hunger contraction was infrequent. Sensations of hunger were present more or less throughout the period, being therefore probably due to the persistently increased tonus. The onset of a period of hunger contraction could usually be foretold by an

increase in the hunger sensation, and as these contractions became more marked, the hunger sensations became more intense. On the last day of starvation a burning sensation referred to the epigastrium was added to that of hunger. The appetite ran practically parallel with the sensation of hunger, and both of these sensations became perceptibly diminished on the fourth or last day of starvation, this diminution being, however, most marked in the sensation of appetite. Indeed, instead of an eagerness for food, there developed on the last day a distinct repugnance or indifference towards it. Accompanying these sensations of hunger and appetite a distinct mental depression and a feeling of weakness were experienced during the latter part of the starvation period.

On partaking of food again the hunger and appetite sensations very rapidly disappeared, and also practically all of the mental depression and a great part of the feeling of weakness. Complete recovery from the latter, however, did not take place until the second or third day after breaking the fast. From this time on both men felt unusually well; indeed they state that their sense of well-being and clearness of mind and their sense of good health and vigor were as greatly improved as they would have been by a month's vacation in the mountains. They further point out that, since others who have starved for longer periods of time unanimously attest the fact that, after the first few days, the sensations of hunger become less pronounced and finally almost disappear, they must have experienced the most distressing period during their four days of starvation. Although the hunger sensation was strong enough to cause some discomfort, it could by no means be called marked pain or suffering, and was at no time of sufficient intensity to interfere seriously with work. Mere starvation can not therefore be designated as acute suffering. It is of further interest to note that during the starvation period a continuous flow of secretion of acid gastric juice was found to occur in the stomach, the presence of this acid probably explaining the acid or burning sensation experienced in the epigastrium on the last days.

CONTROL OF THE HUNGER MECHANISM

The control of the hunger mechanism, like that of any other mechanism in the animal body, may be effected through the nervous system or it may depend on the presence of chemical substances or hormones in the blood. As a matter of fact, it can readily be shown that both those methods of control are operative, and we will now consider briefly some of the facts upon which this conclusion depends.

Although many facts are now known with regard to the nervous con-

trol of the hunger mechanism, it is difficult to piece these together in such a way as to formulate a simple theory which fits in with all the observed facts. We know that the stomach possesses in itself a local nervous mechanism by which, like the heart or intestine, it can automatically perform many of the movements which are exhibited in the intact animal. These local movements may, however, be considerably influenced by impulses transmitted to the stomach along the vagus and splanchnic nerves. We have therefore to seek for evidence indicating the relative importance of the local nervous mechanism in the stomach itself and of the impulses transmitted to this organ by the extrinsic nerves. We must then seek the position of the center which perceives the sensation of hunger.

It will be simplest to consider first the effect of section of the extrinsic nerves in observations made on lower animals. *Section of the splanchnic nerves* increases gastric tonus and augments the gastric hunger contractions. Section of *both vagus nerves,* performed of course below the level of the heart, leaves the stomach in a more or less hypotonic condition. The tonus is not entirely abolished; it varies somewhat from day to day, and may become quite pronounced even though the vagi are cut. In this hypotonic state the hunger contractions are diminished in rate and regularity. Section of *both* the splanchnic and vagus nerves throws the stomach into a permanent hypotonus, except in prolonged starvation, when hunger contractions develop that are usually of great amplitude and with particularly long intervals between the contractions. The general conclusion to be drawn from these experiments is that, although completely isolated from the central nervous system, the stomach still exhibits typical hunger contractions, which must therefore be essentially dependent upon an automatic mechanism in the stomach wall itself. Over this mechanism, extrinsic nerve impulses have merely a regulatory control.

Variations and Inhibitions of the Hunger Contractions

The afferent stimuli that may set up impulses traveling by the extrinsic nerves to the stomach are conveyed by the nerves of sense or are of psychic origin. Stimulation of the gustatory end organs in the mouth, as by chewing palatable food, always causes an inhibition of the tonus and a diminution or disappearance of the hunger contractions. Even the chewing of indifferent substances, such as paraffin, suffices to produce distinct inhibition, unless in a case in which the contraction has passed into a tetanus. It is of interest that swallowing movements, in the absence of any food substance in the mouth, are sufficient to produce a transitory inhibition of the gastric tonus—a receptive relaxation of the

stomach, as it has been aptly called. The diminution in tonus and hunger contractions in these various ways is accompanied by a diminution in the hunger pains.

Afferent nerve stimulation affecting the hunger contractions may also originate in the *stomach mucosa* itself, as has been shown by Carlson on his patient by introducing the various substances to be tested through a tube into the stomach. A glassful of cold water introduced in this way inhibits the tonus and the hunger contractions for from three to five minutes unless these are severe, this inhibition being followed by no augmentation either of the tonus or of contractions. Ice-cold water has a greater effect than water at body temperature. This result is somewhat different from that which most men experience as the result of drinking a glass of cold water.

Weak acids of strengths varying up to that found present in the gastric juice itself—0.5 per cent—cause a marked inhibition of the hunger movements, but this inhibition does not persist until all the acid has escaped from the stomach or been neutralized, which explains why hunger contractions should still occur when an acid secretion is present in the stomach, as in starvation. Normal gastric juice itself produces an inhibition, which is no doubt dependent upon the acid which it contains, and it is probable that, at the same time that it leads to inhibition of the hunger contractions, the acid initiates peristalsis of the pyloric region (see page 453). Weak alkaline solutions have no greater effect on the hunger contractions than an equal volume of water. Weak solutions of local anesthetics, such as phenol or chloretone, are without effect.

With regard to *alcoholic beverages* interesting results were obtained. Wine, beer, brandy, and diluted pure alcohol inhibit both the tonus and the contractions. The duration of this inhibition varies directly with the quantity of the beverage introduced into the stomach and with its alcohol percentage. These observations are apparently not in harmony with the experience of most men that the taking of alcoholic beverages serves to awaken or increase the appetite, the difference being no doubt due to the fact that appetite and hunger contractions of the stomach are not dependent on each other, appetite being, as we have seen, a complex psychic affair, whereas the hunger contractions depend upon a local mechanism in the stomach wall itself.

As the inhibition produced in one or other of these ways passes off, the hunger contractions are resumed at their previous intensity and not in an augmented form. From the promptness of the inhibition, it would appear that the stomach contractions are affected, not reflexly through the central nervous system or by changes in the chemical composition of the blood, but by a direct action on the neuromuscular mechanism

in the stomach walls, and it is important to bear in mind that the inhibitory effects on the stomach contractions of the fundus may proceed quite independently of the changes in the pyloric region that are concerned with the mechanical processes of digestion. After one or both of the extrinsic nerves of the stomach were severed in dogs, a certain degree of inhibition could still be induced by the above methods, indicating that, although section of the extrinsic nerves depresses the inhibitory reflex, it does not abolish it.

Various mitigations of the hunger contractions have been discovered. Smoking has this effect, and compression of the abdomen by tightening the belt also inhibits the contractions provided they are not of marked intensity. Considerable muscular exercise, such as brisk walking or running, causes inhibition, which usually persists until after the exercise is discontinued. When the tonus and contractions return, in this case, they seem to be somewhat more pronounced. Application of cold to the surface of the body—as by placing an ice pack on the abdomen or taking a cold douche, procedures which are well-known to induce increased neuromuscular tonus, in general—causes an inhibition of the gastric tonus and hunger contractions, the degree of which is roughly proportional to the intensity of the stimulation. There is certainly never an increase in the gastric tonus or hunger contractions. If such stimulation is maintained, the inhibitory effects on the stomach gradually diminish, even though the individual be shivering intensely.

With regard to *the nerve centers concerned in these phenomena,* little that is definite is known. The sensory nuclei of the vagus nerve in the medulla must be considered as the primary hunger center, and through this center, not only influences affecting the stomach contractions, but also those associated with the hunger sensations, must be mediated. It would appear from observations on the hunger behavior of decerebrate animals that there can be no hunger center located on the cerebral cortex itself, for such animals exhibit practically the same hunger effects as normal animals. It is interesting to note that, at least in the case of decerebrate pigeons, this hunger behavior entirely disappears on removal of the optic thalami, where important nerve centers having to do with the bodily responses of the animal to hunger impulses would therefore appear to be located. These observations support the suggestion that has been made by several neurologists that the sense of pain is located somewhere in the thalamic region.

Concerning the influence of psychic states, Carlson says that in his own case the hunger contractions became weaker and the intervals between them greater when he was suddenly awakened during his fast and saw two of his friends partaking at his bedside of a "feast of

porterhouse steak with onions, potatoes, and a tomato salad." These results are no doubt due to local inhibition dependent upon the psychic secretion of appetite gastric juice. When no such juice is produced, the sight and smell of good food does not appear to affect materially the hunger contractions of the stomach. No doubt it stimulates the appetite, but that, as we have seen, is a psychic affair.

CHAPTER LV

THE BIOCHEMICAL PROCESSES OF DIGESTION

In a book designed primarily for clinical workers, it would be out of place to enter into details concerning the biochemical processes taking place during the digestive process. There is, however, a certain amount of fundamental knowledge which it is essential that we should consider. In the first place it should be borne in mind that in the digestion of carbohydrates and proteins, various intermediate stages are passed through before the final absorption products are formed. The highly complex molecule of which protein, for example, is composed, is first of all broken down into several smaller but still highly complex molecules, each of which then undergoes further disruption, until ultimately the amino acids are set free. Certain enzymes, such as trypsin, can carry this process from the beginning through the greater part of its course without the assistance of other enzymes, but in the natural process of digestion, as it occurs in the gastrointestinal tract, the different stages of the disruption are controlled by different enzymes. One enzyme prepares the food for action by the next. This interdependence of the actions of the enzymes demands that some provision should be made whereby each enzyme is secreted at the proper time; that is, when the foodstuff has already been prepared for its action by that of its predecessor. Thus, it would be useless after food is taken for the gastric and pancreatic juices to be secreted at the same time. Instead, the gastric juice is secreted first, and the pancreatic only after the food has been prepared for its action. This correlation in function we have already seen to be dependent largely on the action of hormones.

DIGESTION IN THE STOMACH

The gastric juice contains two important digestive agencies: (1) the enzyme, pepsin, and (2) hydrochloric acid. It is particularly in juices secreted in the cardiac end of the stomach that these two substances are found present; towards the pyloric end the hydrochloric acid entirely disappears, and the pepsin content becomes distinctly less.

The Functions of Hydrochloric Acid

The functions of hydrochloric acid may be conveniently divided into physiological and biochemical. The former functions have to do with the control of the movements of the stomach, including the opening of the pyloric sphincter, and, after the chyme has entered the duodenum, with the secretion of pancreatic juice and bile. The biochemical functions are concerned: (1) in assisting the pepsin in the digestion of proteins, (2) in bringing about a certain amount of inversion of disaccharides, and (3) in having an antiseptic action on the stomach contents. Regarding the last mentioned of these functions, it may be said that the chyme, as it is ejected from the stomach, is usually sterile, although it may contain spores and certain bacteria that are protected against the digestive agencies of the stomach. This protection is afforded by an outer covering of a chitinous nature (spores), or, as in the case of the tubercle bacillus, by a covering of waxlike material. It is believed that persons with strictly normal digestion are much less liable to infection by such bacteria, as those of typhoid and cholera, than persons with less active gastric secretion. When the acid of the gastric juice falls below the level at which it develops an antiseptic action, various bacteria and yeasts grow in the stomach contents, producing by the resulting fermentation irritating organic acids and gases. It is under these conditions that yeasts, sarcinæ, and lactic and butyric acid bacilli find in the gastric contents a suitable nidus on which to grow.

The Amount of Acid

It has long been known that considerable variations in the amount of hydrochloric acid in the gastric juice are associated with symptoms of indigestion. On this account a more or less elaborate technic has been developed for the purpose of determining the amount of hydrochloric acid in the gastric contents.* There are three things in connection with this activity that we may measure: (1) the total titrable hydrochloric acid; (2) the free hydrochloric acid; and (3) the actual hydrogen-ion concentration. The determination of the total available acids is made by titrating a measured quantity of gastric juice against a standard alkali, using phenolphthalein as an indicator. By this method about 75 c.c. of decinormal alkali solution are required to neutralize 100 c.c. of normal gastric juice. The determination of the free hydrochloric acid is made by using special indicators, such as those of Günzberg and Töpfer, which change color at a hydrogen-ion concentration of about 10^{-5} (see page 27). To produce this hydrogen-ion concentration, a con-

*The methods can be found in any volume on clinical diagnosis.

siderable quantity—0.05 per cent or more—of an organic acid is necessary, whereas it requires only a trace of hydrochloric acid. Normal human gastric juice, when titrated with one of these indicators, gives a figure which corresponds to about 0.03 N. hydrochloric acid (see page 22). For the accurate determination of the hydrogen-ion concentration, it is necessary to use the gas-chain method (see page 29).

When gastric juice is collected through a fistula from an empty stomach, very little difference will be found between the free hydrochloric acid and the total acid; that is, between the results obtained by the second and the first of the methods described above. This is because in such juice there is no organic matter capable of combining with the hydrochloric acid, and there are no other acids, such as lactic or butyric, which might be produced by fermentative processes. The difference between the two titrations, however, becomes quite marked when protein food is undergoing digestion in the stomach, because at its different stages of digestion protein combines with increasing quantities of the hydrochloric acid. The pathologic condition in which there is most definitely a diminution of the hydrochloric acid is cancer, either of the stomach itself or occasionally of some other part of the body. An increase is particularly marked in ulcer of the stomach. A considerable variation in hydrochloric acid may however be the result merely of functional (neurotic) conditions.

The Source of the Acid

A question that has puzzled physiologists for many years concerns *the mechanism by which hydrochloric acid is secreted.* The percentage of hydrochloric acid in the gastric juice is considerably above that at which any animal cells can live, and yet this acid is secreted by the lining membrane of the stomach, its source being, of course, the sodium chloride of the blood plasma. How then do the cells of the gastric glands bring about the separation of this powerful acid from the perfectly neutral blood plasma? In the first place, it is significant that the mucous membrane of the stomach contains a higher percentage of chlorine than the average of other organs and tissues, indicating that it has the power of abstracting chlorine from the blood. The excess of chlorine in the mucosa must, moreover, be but a very small fraction of that actually secreted into the the gastric juice. The chlorine content of the mucosa of the cardiac end is considerably greater than that of the pyloric. These facts indicate that chlorine is attracted by the gastric cells, but they throw no light on the question as to where the hydrochloric acid is really formed. Is it in the cells, or only in the lumen of the gland tubes? That is to say, is it formed before or after the gastric

juice has been secreted from the cells? After intravenous injection of solutions of potassium ferrocyanide and some inert salt of iron, such as one of the scale preparations, examination of the gastric glands has shown that the prussian blue reaction, which requires the presence of free mineral acid, is most pronounced in certain of the parietal cells. A considerable amount of the precipitate is, however, also visible in the lumen of the glands and in the stomach itself. Certain observers affirm that, although some of the parietal cells may take the stain, the vast majority of them do not do so; and, moreover, that cells incapable of forming hydrochloric acid (e. g., of the liver) may also become stained, and that the precipitation may occur in the blood and lymph.

The confusion in the results by these methods prompted A. B. Macallum[14] and Miss M. P. Fitzgerald to investigate the distribution of the chlorine in the cells by a microchemical method, in which the chlorides were precipitated with silver nitrate and the silver chloride then reduced by exposing the section to light. It was found that both kinds of gastric-gland cell, chief and parietal, but particularly the parietal, gave the chloride reaction. Using as a stain a substance (cyaninine) which reacts blue with acid and red with alkali, Harvey and Bensley,[15] however, aver that the secretion of the glands is practically neutral until the foveola is reached, where the stain becomes blue, indicating an acid reaction. This seems to show that the acid is not really secreted by the cells of the gastric gland, but is formed after secretion.

According to the latter investigators, the chlorine is secreted by the cells into the fovea as some weak chloride, such as ammonium chloride, or it may be as an ester. Shortly after its secretion this weak chloride undergoes a hydrolytic or other dissociation, during which free hydrochloric acid is liberated and ammonia or some other weak base set free. Of these two products of the reaction the weak base is reabsorbed by the gland cells, but the hydrochloric acid is left behind because the cells are impervious to it. Indirect evidence in support of this view is afforded by certain other instances in which hydrochloric acid is produced by the action of cells; thus, the mold Penicillium glaucum when it is grown in a medium containing ammonium chloride absorbs the ammonia but leaves the hydrochloric acid. The high penetrating power of the ammonia ion in practically all cells, and the fact that the mucosa of the stomach contains a higher percentage of ammonia than any other tissue in the body, must also be considered as circumstantial evidence in favor of this view.

Whatever be the mechanism by which hydrochloric acid is produced, there is no doubt that the epithelium is impenetrable to it. When the vitality of the epithelium becomes lowered, as in anemia or after partial

occlusion of the arteries, the acid may penetrate the cells and cause digestion of the stomach walls. Hyperacidity may on this account become dangerous, as it lowers the resistance of the cell.

The digestive action of hydrochloric acid is closely linked with that of pepsin, with which it will, therefore, be considered.

The Action of Pepsin

It is commonly believed that before its secretion pepsin exists in the cells of the gastric glands as zymogen granules. The chief evidence for this belief appears to be that after considerable activity the amount of zymogen granules in the gland cells is found to be decidedly diminished. By such an hypothesis it is easy to explain certain interesting results concerning the effect of weak alkali on the activities of extracts of the mucous membrane of the stomach. When the mucous membrane is extracted with weak acids, the extract is very active proteolytically. If this so-called pepsin solution be made faintly alkaline, or even only neutralized, and again made acid, it will be found to have lost much, if not all, of its activity. On the other hand, an aqueous extract may be rendered slightly alkaline for a short time and still display its digestive activity on subsequent acidification. The extract made with water is therefore much more resistant toward alkali than that made with weak acid, and the difference is explained on the supposition that the watery extract contains pepsinogen, whereas the acid extract contains pepsin.

It is believed that there are several varieties of pepsin, because the optimum concentration of acid in which pepsin derived from the stomachs of different animals acts is not always the same. Pepsin of the dog, for example, acts best in a hydrogen-ion concentration corresponding to that of a 0.05 N. hydrochloric acid solution, whereas that of the human stomach works best at a concentration of 0.03 N. Different pepsin solutions also show a difference with regard to the optimum temperature at which they act, and with regard to the nature of the protein which they most readily attack. Thus, the pepsin of a calf's stomach digests casein very rapidly, but coagulated egg white only slowly, whereas the pepsin of the pig's stomach acts on both these proteins at about the same rate.

It is well known that the activity of pepsin can proceed only in the presence of acids, but this action of acids does not appear to depend on the hydrogen-ion concentration alone, for when equal quantities of the same pepsin are mixed with quantities of different acids so that the hydrogen-ion concentration of the mixtures is uniform, it is found that digestion proceeds most rapidly with hydrochloric acid and least rapidly with sulphuric acid. The SO_4 ion seems, therefore, to be unfavorable

for peptic activities. The acid seems to combine with the protein before the pepsin attacks the latter; for, if we first combine the protein with acid and then wash away all traces of free acid, the protein can be digested in a neutral pepsin solution without the liberation of any free acid.

There is evidence to show that pepsin itself also becomes combined with the protein during the digestive process. If a piece of protein such as fibrin be immersed in a solution of pepsin and then taken out and washed thoroughly to get rid of all adherent pepsin, it will be found, on placing it in a hydrochloric acid solution of the proper strength, that peptic digestion proceeds. Advantage may be taken of this fact to separate pepsin from a solution, but the best protein to use for this purpose is not fibrin but elastin. By such a method it has, for example, been shown that there is some pepsin in the intestinal contents, proving thus that when the chyme passes into the intestine, the pepsin is not, as used to be thought, immediately killed by the proteolytic enzyme.

Products of Peptic Digestion

With regard to the products of gastric digestion, little can be said here. The first product is a metaprotein known as acid albumin or syntonin. It is precipitated from the digestion mixture by neutralization. The next product is known as primary proteose, being precipitated by half saturation with ammonium sulphate. The third product is secondary proteose, produced by complete saturation with the above reagent; and after all these bodies have been separated out, there remains in solution the fourth product—peptone—which among other things is characterized by the fact that with the biuret test it gives not a violet but a rose-pink color.

It has often been claimed that along with these products a certain amount of free amino acids may also appear in a peptic digestive mixture. This, however, may be due to the action of erepsin, which is usually present in pepsin preparations. It is important to note that the term *proteose* is a general one, and that there are probably many varieties of this substance, differing from one another according to the protein from which they are derived.

The change produced by pepsin and hydrochloric acid is of the nature of an hydrolysis, for it has been found that the amount of hydrogen and oxygen in the digestive products is greater than that in the original protein. It is by a similar process of hydrolysis that the other proteolytic enzymes, such as pancreatin and erepsin, operate, but this does not imply that the exact grouping that is split apart by the hydrolytic proc-

ess is the same for each of these enzymes. Indeed, there is considerable evidence that pepsin does not, like the other enzymes, break up the long chain of amino acids that are linked together to compose the polypeptides, but that it only splits the big molecule of albumin or globulin into several large groups, each of which is composed of long amino-acid chains. Its action appears to be analogous with that of amylase on starch, by which, it will be remembered, the big polysaccharide molecule is split into smaller polysaccharide molecules, which then become attacked by the dextrinase and split into disaccharide molecules (see page 656). The evidence in support of this view is: (1) that pepsin is unable to digest polypeptides, and (2) that it is able to digest certain proteins upon which erepsin (see page 490) has no action.

The hydrolytic splitting of large into smaller protein molecules, like that by which the chains of amino acids in the polypeptides are subsequently broken up, consists in a breaking of amino-carboxyl linkings (NHCO) (see page 598), with the consequent liberation of a large number of unattached amino groups. The number of these free amino groups can be determined quantitatively by the formaldehyde titration method of Sörensen.* By this method it can be shown that from the very start of peptic digestion the number of free amino groups increases, and *pari passu* the power of the digestive products to combine with free hydrochloric acid. Indeed, when the experiments are done quantitatively and the digestion allowed to proceed for a considerable time, the increase in the formol titration is practically equal to the decrease in the free acids as determined by the Günsberg reagent.

The rate of peptic digestion is usually estimated by the law of Schütz and Borissow, according to which the amount of coagulated albumin that is digested in a Mett's tube is proportional to the square root of the amount of pepsin.†

The pepsin which leaves the stomach in the chyme is not all destroyed in the intestine, as was at one time believed to be the case, for, as we have seen above, some pepsin can be detected in the gastrointestinal contents. A part of the pepsin may be absorbed into the blood and carried back to the gastric glands to be used again. This would account for the presence of antipepsin in the blood, and also for the presence of pepsin in the urine. It is probable, however, that most of the pepsin is destroyed after it enters the intestine.

*In this method the basic character of the amino acids is destroyed by the formaldehyde, so that a higher degree of acidity develops in the mixture. By determining the increased acidity by titration with alkali, an estimate is obtained of the number of amino groups. (See page 599.)

†The amount of coagulated egg albumin digested is ascertained by measuring the length digested away from the end of a column of coagulated egg white contained in a glass tube (Mett's method). (See Cobb, P. W.: Am. Jour. Physiol., 1905, xiii, 448.)

Clotting of Milk in the Stomach

Besides its power of digesting protein, the gastric juice is also endowed with the property of clotting milk. This action is commonly attributed to the presence of another enzyme besides pepsin, namely, ***rennin;*** but in recent years considerable controversy has raged around the question as to whether pepsin and rennin are not the same thing. One strong argument in favor of this view is that all digestive juices that are capable of digesting protein can also clot milk. In any case, when gastric juice acts on milk, it splits the casein* of the milk into two portions, one of which, called paracasein, immediately combines with calcium to form an insoluble colloidal compound, which is precipitated and, by entangling the fat of the milk, forms the clot; the other protein remains in solution and is known as whey albumose. From studies on molecular weight it is believed that the paracasein is produced from casein by the splitting of the molecule of the latter into two, from which it would appear that the action of this enzyme is nothing more than the first stage in the hydrolysis of the casein molecule. The whey albumose, according to this view, is a by-product.

There are many investigators, however, who believe that rennin and pepsin are not identical, since an infusion of the stomach of a calf has a powerful clotting action on milk but a very weak digestive one on egg white, whereas a similar infusion from the stomach of a pig shows exactly the reverse properties. This question is one of so controversial a nature that it would be out of place to go into it further here. It should be pointed out, however, that, when the gastric contents are acid in reaction, milk will become clotted by the action of the acid itself quite independently of any pepsin or rennin the juice may contain. This acid clotting of milk is probably of a different chemical nature from that produced by the enzymes.

On other foodstuffs than proteins the action of the gastric juice is relatively unimportant, although polysaccharides may be considerably broken down in the cardiac end of the stomach on account of the action of swallowed saliva (see page 454), and disaccharides, as we have seen, may become split by the hydrolyzing effect of the hydrogen ion. Fat digestion also takes place in the stomach when the fat is taken in an emulsified condition, as in milk and egg yolk, but not when in masses, as in meat or butter. This action is due to the presence of a fat-splitting enzyme, or lipase, in the gastric juice.

---

*In the above nomenclature casein is the same as caseinogen, and paracasein the same as casein, of the English physiologists.

CHAPTER LVI

THE BIOCHEMICAL PROCESSES OF DIGESTION (Cont'd)

DIGESTION IN THE INTESTINES

The further changes which the half-digested foodstuffs in the chyme undergo in the intestinal canal depend on the enzymes present in the secretion of the various glands and on the presence of bacteria. The most important of the digestive juices are the pancreatic juice and bile. The latter, however, does not contain any enzyme, its influence on digestion being entirely adjuvant.

Pancreatic Digestion

When we were considering the mechanism of secretion of the pancreatic juice, we saw that the juice produced by the action of secretin on the gland cells does not contain any active proteolytic enzyme, although it contains one capable of acting on polysaccharides and another, on fat.

The Action of Trypsin

When this juice is mixed with the secretion of the duodenum or of the upper part of the small intestine, it immediately develops powerful proteolytic power. The same result may also be obtained by mixing it with an extract of the mucous membrane of the duodenum made with dilute bicarbonate solution. A very small amount of the extract is capable of increasing the digestive activity of a very considerable quantity of pancreatic juice, showing that the action depends on the presence of an enzyme which has been called *enterokinase*. This influence of the intestinal secretion is readily destroyed by heating.

Large quantities of alkali are contained in the pancreatic juice and bile, so that in the upper reaches of the intestine the acidity of the chyme is practically neutralized. A little lower down, however, an acid reaction may again develop (see page 505). On account of these facts it has been concluded that the activity of trypsin is most rapid in the presence of a slight excess of hydroxyl ions; i. e., in a weakly alkaline solution. It is interesting to note that, as a result of the great secretion of alkali by the pancreas, extracts of this organ after death show a very high degree of acidity in comparison with extracts from other organs

and tissues. It has also recently been shown that the activity of trypsin does not depend on the presence of free hydroxyl ions, but that it may proceed in the presence of free acid, even up to a strength of $C_H = 1.5$. If pepsin is present together with trypsin in a distinctly acid solution, the pepsin seems to destroy the trypsin, unless the mixture contains a considerable quantity of protein, when the tryptic activity may persist even for several hours. A practical conclusion that we may draw from these results is to the effect that preparations of trypsin—the so-called *pancreatin*, for example—if given with the food, may pass in an active condition into the duodenum, where, in the more favorable environment created by the neutralization of the excess of acid, it will develop its proteolytic power. The therapeutic administration of pancreatin is, therefore, justified (Long[16]).

The activated trypsin acts on proteins in very much the same way as pepsin, except that the decomposition of the peptone and proteoses into polypeptides is the chief feature of the process. Thus, after tryptic digestion has proceeded for some time, only a trace of primary proteoses but considerable quantities of leucine, tyrosine and other amino acids will be found present. Some investigators believe that the thorough nature of the digestive action of activated pancreatic juice may depend on its also containing erepsin, an enzyme which we shall see to be present in considerable amount in the mucous membrane of the intestine and other tissues, and whose particular function is to split polypeptides into the amino acids. From the autolytic digestion which takes place in organs kept in a sterile condition after death, tryptic digestion differs in that it produces only small quantities of ammonia. The large quantities of ammonia produced in autolytic digestion no doubt have a relationship to the acids simultaneously set free during this process.

In the products of tryptic digestion it is usually found that, although there has been considerable splitting of the protein into amino acids, there are still a good many amino-carboxyl (NHCO) linkages left unbroken, indicating that certain polypeptides are left intact in the mixture. To split the polypeptides requires the aid of the *erepsin*, which is present in the mucous membrane of the intestine. Interesting investigations have been made on the exact degree to which trypsin-enterokinase can split up the various known polypeptides. This seems to depend on the structure of the polypeptide molecule and on the number of amino acids present in the chain. For example, analylglycine, but not glycylalanine is hydrolyzed, although both contain the same amino acids but linked together in a different way; and tetraglycylglycine, which contains five glycine radicles, is hydrolyzed, whereas diglycylglycine, which contains only three, is not.

The importance of the presence of *erepsin* in the mucous membrane of the intestine is that it serves as a barrier to the passage of any unsplit amino acids from the intestinal contents into the blood. It insures the breaking up of the protein molecule into its ultimate units before absorption. The further fate of the absorbed amino acids will be considered under the subject of protein metabolism.

The Action of Lipase

Neutral fat is decomposed into fatty acids and glycerine by the *lipase* present in the pancreatic juice. This enzyme may also be extracted from the glands by means of 60 per cent alcohol. Its action is remarkably accelerated by the presence of bile, and considerably depressed by inorganic salts. It is also very dependent on the degree of alkalinity, the optimum being a hydrogen-ion concentration of $H \times 10^{-8}$. The favoring action of bile is undoubtedly owing to the bile salts (see page 493), and it is probable that this action is dependent upon the influence which these have in lowering surface tension and therefore bringing about a more intimate contact between fat and water.

The Action of Amylopsin

The action of pancreatic juice on carbohydrates depends on the amylolytic enzyme called *amylopsin*. In animals having no active ptyalin in the saliva, amylopsin serves as the only diastatic enzyme concerned in the digestive process. In any case, at least for the first stages of the disruption of the starch molecule—that is, its conversion into dextrines—amylopsin is a more powerful enzyme than ptyalin. It does not appear to be so efficient as ptyalin in the final stages of the hydrolysis, for it does not produce so much reducing sugar as ptyalin does. Indeed extracts of pancreas will sometimes convert starch into soluble starch and dextrine with great speed, but produce scarcely any reducing sugar. On this account it is believed by many investigators that there are at least two distinct and separate enzymes in amylopsin and also perhaps in ptyalin, one a true amylase, which converts starch into dextrine, and the other a dextrinase, which converts dextrine into maltose. In the case of both ptyalin and amylopsin digestion proceeds best in a very weak acid reaction. Amylopsin, as it is secreted in the pancreatic juice, is fully activated; bile, apart from the alkali which it contains, having no influence on its digestive power.

Besides amylopsin the pancreatic juice also contains *maltase,* and in the case of young animals or of those that take milk with their food throughout their lives, lactase also. After the suckling animal has dis-

continued taking milk, the lactase disappears from the pancreatic juice. Attempts have been made to bring it back by feeding the adult upon milk, but without success. Occasionally the pancreatic juice also contains *invertase*.

The Bile

Associated with the pancreatic juice in all its functions is the bile. When this fluid is prevented from entering the intestine, the digestive process becomes very imperfect, the absorption of fat being particularly interfered with (see page 691). Bile is also an excretory product, and its *composition* therefore is much more complex than that of the other digestive fluids. This varies very much, however, according to the method of collection. Bile from the gall bladder after death contains much more solid material, particularly bile salts and mucin, than that collected from a fistula of the bile duct or gall bladder during life. These differences will be evident from the accompanying table.

100 parts contain—	Bile from Gall bladder	Bile from Fistula
Water	86	97
Solids	14	3
Organic salts (bile salts)	9	0.9-1-8
Mucin and bile pigment	3	0.5
Cholesterol	0.2	0.06-0.16
Lecithin and fat	0.5-1.0	0.02-0.09
Inorganic salts	0.8	0.7-0.8

In general it may be said that bile obtained from a fistula in man contains only about 3 per cent of total solids, of which from one-fourth to one-half are inorganic, whereas bile from the gall bladder contains 10 to 20 per cent of total solids, of which only about one-twentieth are inorganic. The chief cause for this difference appears to be that when the bile goes to the intestine, a considerable proportion of its bile salts is reabsorbed into the portal blood and reexcreted by the liver. Some of the difference may also be caused by the fact that absorption of water takes place from the gall bladder, and that mucin and possibly cholesterol are secreted by this organ. These striking differences between fistula and gall-bladder bile are observed only when the common bile duct is occluded. If the bladder fistula is made with the common duct left open, some of the bile gains entry to the duodenum and therefore becomes reexcreted. It is well known that a fistula of the gall bladder in man after a time closes up and the bile again takes its usual course along the bile duct into the duodenum.

Interesting observations have been collected on the amount of the secretion from a fistula both in man and in the lower animals. In man it is commonly stated that about 500 c.c. of bile are secreted daily, the amount varying considerably during the different hours of the day. The secretion of bile is greatly reduced by hemorrhage. It is greater on a meat diet than on one of carbohydrates. It is reduced during starvation, but continues to be secreted up to the moment of death.

Functions of Bile

One of the main functions of the bile salts is that they greatly assist, not only in the digestion, but also in the absorption of fats. When bile is excluded from the intestine, the feces are loaded with fatty acids which have been split off partly by the now less effective lipase and partly by the action of bacteria. The fatty acid thus liberated in the absence of bile salts is not absorbed, because the bile salts serve as the carriers of fatty acids into the epithelial cells and lacteals. They combine with the fatty acids, probably by forming some chemical compounds, in which they carry them into the endothelial cells where the compounds become disrupted, the fatty acid combining with glycerine to again form neutral fat and the bile salts being carried to the liver and reexcreted. The influence of bile salts in assisting the action of lipase is probably due to a lowering of the surface tension, thus bringing water and fat into closer union. This accelerating influence has also been demonstrated when synthetic bile salts have been used, showing clearly that it is really these and not any other constituent of the bile that are responsible for its accelerating influence.

Bile also functionates as a *regulator of intestinal putrefaction.* This it does apparently because of its slight laxative properties, by which the intestinal contents are expelled before the bacteria have grown to any great extent in them. Bile itself is a favorable culture medium for certain bacteria, so that it can have no antiseptic action. Its assistance in the action of trypsin and amylopsin depends very largely upon the alkali which it contains.

As an *excretory vehicle* bile is important, because it possesses the power of dissolving cholesterol. Toxins and metallic poisons of various kinds are also excreted in it.

Although not directly concerned with the digestive function, it will be convenient to say something here concerning the chemical nature and derivation of the various biliary constituents.

THE CHEMISTRY OF BILE

The Bile Salts

In most animals the bile salts consist of the sodium salts of glycocholic and taurocholic acids. Each of these acids is composed of a part called cholic acid which is more or less related to cholesterol, and of glycine (CH_2NH_2COOH amino-acetic acid) or taurine ($C_2H_7NSO_3$), a derivative of cysteine, which is α-amino-β-thiopropionic acid ($CH_2HS.CHNH_2.COOH$). The exact form of cholic acid varies in different animals, that of the pig, for example, being different from that of man. Bile salts are an exclusive product of liver metabolism; i. e., they are not formed in any other part of the animal body. They give a very sensitive color reaction known as Pettenkofer's, which however is not specific of bile acids, since it is also given by oleic acid and by many aromatic substances and alcohols. It must be remembered that the part of the bile salts that is characteristic of the liver is the cholic acid, the taurine and glycine being present in other tissues and organs.

When cholic acid is given to animals mixed with the food, the amount of taurocholic acid excreted with the bile is increased, indicating that there must be a store of taurine available in the organism. This store can not, however, be large, for if the feeding with cholic acid is repeated several times, it will be found that the taurocholic acid diminishes and glycocholic acid takes its place; and this increased excretion of glycocholic acid goes on just as long as cholic acid is fed. The reserve of taurine in the animal body appears therefore to be limited, although it is used in preference to glycine when there is an excess of cholic acid to be neutralized. On the other hand, the store of glycine seems to be inexhaustible. That there is no reserve of cholic acid itself in the body is indicated by the fact that no increase in taurocholic acid excretion by the bile results when cystine, the mother substance of taurine, is given with the food. If both taurine and cholic acid be fed, however, the excretion of taurocholic acid increases.

The relative amounts of taurocholic and glycocholic acids in the bile of different animals differ considerably. Human bile contains relatively a small amount of taurocholic acid; on the other hand, the bile of the dog contains a large excess of it.

Cholesterol

In human bile the percentage of this important substance is not high (1.6 parts per 1000), but it is of great clinical importance because of the fact that it may separate out as a precipitate forming *gallstones*. The

percentage of cholesterol in these varies from 20 to 90; the remainder being organic material such as epithelial cells, inorganic salts, pigment, etc. The origin of cholesterol is partly endogenous and partly exogenous. In the former case it comes from the envelope of red blood corpuscles and from the nervous tissues, where it is present in considerable amount. The latter source is, of course, the food. The increase in cholesterol esters in the blood after feeding with food rich in this substance has been shown, particularly in rabbits.

That the bile should be the pathway through which cholesterol is excreted depends no doubt on the fact that it contains bile salts, which along with their other properties have a remarkable solvent action on cholesterol. This solvent property depends on the cholic acid part of the bile salts, which, as already remarked, is chemically very closely related to cholesterol; indeed, the relationship is so close that some have suggested that cholic acid is derived from cholesterol. This would mean that the cholesterol of blood is excreted in two ways, as cholesterol and as cholic acid. Other observers, however maintain that the cholesterol is excreted mainly by the lining membrane of the gall bladder, and that this explains why gall-bladder bile contains more of it than fistula bile. This evidence is, however, not very strong, for the greater excretion of cholesterol under conditions where the circulation of bile is going on may be explained as due to the presence of bile salts, which serve to carry the cholesterol out of the blood.

Many problems remain to be elucidated in connection with the metabolic history of cholesterol. That some of it is absorbed when cholesterol is contained in the food might seem to indicate that its source is entirely exogenous. Against this view, however, stand two facts: (1) that the cholesterol in the feces of herbivorous animals is of the same variety as that present in those of carnivorous animals and not the phytosterol which is present in plants; and (2) that the universal presence of cholesterol in cells indicates that it must be manufactured there.

The Bile Pigments

The pigments of bile are *bilirubin* and *biliverdin*. The latter is produced from the former by oxidation. If the oxidation be carried a stage further, a blue pigment called bilicyanin is formed. This process of oxidation can be observed in the ring test for bile pigment with fuming nitric acid. When bilirubin is reduced, urobilin, one of the pigments in urine, is formed. Bilirubin must therefore be considered as the mother substance of all these pigments, and it is of interest in connection with its derivation to know that it has the same formula

as iron-free hematin or hematoporphyrin, which is produced by treating hemoglobin with concentrated sulphuric acid.

Chemical investigation has shown that bilirubin is built up from substituted pyrrols, probably four such being contained in the molecule. The pyrrol group is also present in indole and tryptophane, and consists of four carbon atoms and an NH group linked together as a ring (see page 604). Similar pyrrol derivatives can be produced by decomposing chlorophyl, the green coloring matter of plants. It is important to remember that bilirubin is acid in nature, and, therefore, can combine with alkalies to form salts. The relative amounts of bilirubin and biliverdin vary in the bile of different animals.

When these pigments enter the intestine they are reduced to urobilin, part of which passes out with the feces, another part being absorbed into the blood and excreted in the urine. Part of that excreted in the urine exists, however, as a so-called chromogen named *urobilinogen.* The urobilinogen is converted into urobilin by the action of oxygen.

The method by which urobilin is produced from blood pigment has been studied by histologic examination of the liver particularly of birds and amphibia, in which destruction of blood pigment goes on rapidly. Increased destruction of blood pigment can be induced by poisoning with certain substances such as arseniureted hydrogen. From such studies it is usually believed that the bile pigments are a peculiar product of hepatic activity, being produced from blood pigments that are derived from erythrocytes which have been broken down either in the liver itself or in some other viscus (e. g., the spleen). Whipple and Hooper[20] have brought forward seemingly incontrovertible evidence against such a view. They have found, for example, that the bile pigments are formed just as readily in animals in which the circulation of the liver was greatly curtailed by anastomosing the portal vein with the vena cava (Eck fistula) as in normal animals. Even when the circulation was limited to the anterior end of the animal (head and thorax) bile pigment appeared in the blood when hemolyzed erythrocytes were injected, and it was also formed when hemoglobin was placed in the pleural and peritoneal cavities. The endothelial cells of the blood vessels and elsewhere can evidently form the pigments, at least when the liver is absent. When such a process occurs under normal conditions, it is quite probable that the liver acts merely as an excretory organ for the pigments in the same way as the kidney does for urea. Possessed of endothelial cells, the liver might itself also produce some of the pigments, but no more than other organs with a similar number of those cells.

Even the derivation of bile pigments from hemoglobin is called in question, for the same workers have observed that, whereas the excre-

tion of pigment from a biliary fistula is remarkably constant in a dog fed on a fixed mixed diet, it became increased, sometimes by 100 per cent, when the diet was changed to one of carbohydrates, and depressed on a diet of meat. The question arises as to whether, after all, the bile pigments are really derived from broken-down hemoglobin. May they not be manufactured *de novo* out of other materials?

Whipple and Hooper have also shown that bile is a most important secretion, for dogs rarely survive on an ordinary diet if bile is permanently prevented from entering the intestine. Intestinal symptoms soon supervene, and become progressively more severe until the death of the animal. Feeding with bile does not relieve the condition, but feeding with cooked liver seems to have a beneficial effect.

After extravasation of blood in the subcutaneous tissues, as in a bruise, for example, a decomposition of hemoglobin proceeds quite like that occurring in the liver, and leads to the production of blue and brown and green pigments like those of the bile. When hemolysis is produced, as by inhalation of arseniureted hydrogen or the injection of inorganic or biological hemolysins, there is an immediate increase in the amount of bile pigment in the bile. Even the injection of hemoglobin solutions has this effect. Under these conditions of hemolysis, besides an increase in urobilin, there may be considerable quantities of hemoglobin secreted in the urine.

Bile salts and pigments usually accompany each other when anything occurs to interfere with the free secretion of bile. For example, after ligation of the bile duct both bile pigments and bile salts accumulate in the blood, in the serum of which they may be recognized by the ordinary chemical tests in from four to six hours after the operation. If the accumulation be allowed to proceed further, the bile pigments become deposited in the tissues, giving them the peculiar yellowish appearance known as jaundice. Under these conditions the bile salts and pigments also appear in the urine. The accumulation of bile salts in the body affects certain physiologic processes; for one thing, it causes a great lengthening in the clotting time of the blood.

If the blood supply to the liver is interrupted by ligation of the portal vein and hepatic artery at the same time that the bile ducts are occluded, not a trace either of bile salts or of bile pigment appears in the blood during the six to eighteen hours that the animals survive the operation.

The amount of obstruction of the bile duct necessary to produce these symptoms is very slight, since bile is secreted at a very low pressure. Even a clot of mucus or a swollen condition of the mucous membrane of the duct is sufficient to produce obstruction. In the discharge of bile from the gall bladder into the duodenum it is claimed by Meltzer[21] that a

reciprocal relationship exists between the contraction of the bladder musculature and the relaxation of the muscular fibers surrounding the duct in the duodenum. If this reciprocal innervation fails to operate properly, discharge of bile into the duodenum may become obstructed so that a certain amount passes back into the blood, as in cases of bile-duct obstruction.

Bile also contains a certain amount of *lecithin and other phospholipins.* The amount varies considerably in the bile of different animals, even in animals of the same species. It is probably derived, as already mentioned, like the cholesterol, from the breaking-down of red blood corpuscles that goes on in the liver. It is no doubt digested by the ferments of the intestinal tract, the liberated cholin, since it is toxic if absorbed, being further attacked by bacteria so as to become converted into certain substances of a nontoxic nature.

CHAPTER LVII

BACTERIAL DIGESTION IN THE INTESTINE

On an average diet, in twenty-four hours the feces of man weigh about 100 grams, or after drying, about 20 grams. About one-fourth of the dry matter consists of the bodies of bacteria. If plated out by the ordinary bacteriologic methods, however, it will be found that only a small proportion of these bacteria are living. The greater number have been destroyed, probably by the action of the mucin in the large intestine. The nitrogen content of the feces amounts to about 1.5 grams a day, of which about one-half is bacterial nitrogen. If the diet contains large quantities of cellulose material, as in green vegetable food and fruit, the mass of feces as well as the bacterial content may be considerably greater.

The foregoing facts indicate that very extensive bacteriologic processes must be going on all the time in the intestinal contents, and the question arises as to whether such action is beneficial or otherwise to the animal economy. To answer this question interesting observations have been made on the growth and well-being of animals excised from the uterus under strictly sterile conditions and maintained thereafter on sterile food. Such observations made on guinea pigs have shown that the animals thrive and grow perfectly for a considerable time. Experiments carried out on chicks have not, however, yielded similar results. Chicks hatched out from the egg under strictly sterile conditions and then fed on sterile grain, do not thrive, but do so if with the grain is mixed a certain amount of fowl excrement. These experiments, apparently contradictory in their results, show that for certain groups of animals bacteria are required, but not for others.

The difference is probably dependent on the nature of the foods. It will be remembered that the size of the large intestine varies considerably according to the nature of the diet (see page 463). Animals taking great quantities of cellulose foodstuffs have very large ceca and very long large intestines; whereas those which, like the cat, live practically entirely on cellulose-free food, have a rudimentary large intestine. The size of the lower intestine is obviously dependent on the presence or absence of cellulose in the food. It will be remembered also that the forward movement of the contents of the large intestine is very slow; indeed, special provision is made, by the presence of the so-called anti-

peristaltic wave, to delay its movement. This suggests that an important digestive process must be proceeding in this part of the gut. In these ways conditions become established in the cecum for the active operation of bacteria. They attack the cellulose, and liberate the more digestible foodstuffs contained in the vegetable cells, also producing out of the cellulose itself materials of nutritive value. The acids that are also produced by this process are neutralized by the carbonates secreted by the mucosa.

In certain herbivorous animals—the ruminants—this process in the cecum is not relatively of such importance, because it takes place in the paunch. The animals swallow the food and it mixes in this part of the stomach with the saliva, so that bacteria and ferments contained in it, called cytases, attack the cellulose, liberating the more easily digested foodstuffs inclosed within the cell walls. As this process goes on acids accumulate in the digestive mixture. The food is then returned to the mouth, chewed over again, and swallowed again into the main stomach, where it is digested. The aid which bacteria render to digestion depends therefore on the nature of the diet. Man, being omnivorous, stands midway between the two groups of animals discussed above. Although the cellulose contained in his food is not itself sufficiently digested to furnish nutriment, yet it is so far acted upon as to permit the rupture of the cell, the contents of which are then digested. The cellulose is, however, of value in furnishing bulk to the intestinal contents—"intestinal ballast," it is sometimes called.

In the small intestine in man there are bacteria capable of acting on carbohydrates and producing from them organic acids, such as lactic, acetic, etc. So long as a sufficiency of carbohydrate exists to encourage the action of these bacteria, others having an action on protein do not seem to thrive. It may be that this is to be accounted for partly by the production of acid substances by the carbohydrate fermentation, and partly by the fact that, as soon as the protein molecule is broken down by the digestive enzymes, its building-stone amino acids are absorbed. There are probably also bacteria in the small intestine capable of splitting fat into fatty acid and glycerine, but practically nothing is known of their action. In the large intestine of man, along with the cellulose-digesting bacteria already mentioned, protein-digesting bacteria are also present. These bacteria belong to the class, Bacillus coli communis, the various members of which are known as faculative anaerobes because they can grow in the presence or absence of oxygen.

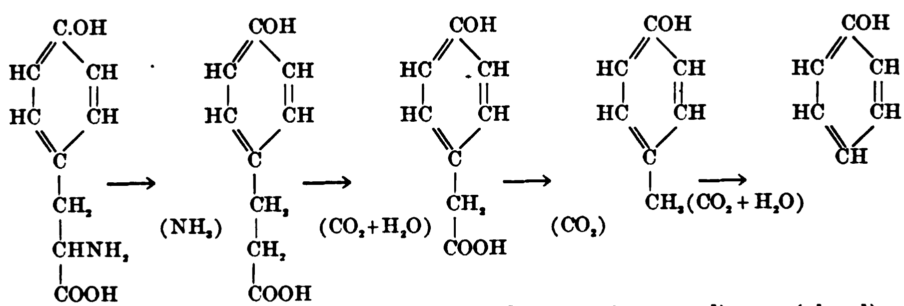
If bacterial growth is excessive or there is an insufficiency of carbohydrates in the small intestine, the bacteria attack the amino acids produced by the digestive enzymes and decompose them into products that may be toxic if absorbed into the blood.

Bacterial Digestion of Protein

From a pathologic standpoint, the most important action of bacteria is that which takes place on protein. Under anaerobic conditions the intestinal bacteria have in general the power of splitting off the amino group whereas under aerobic conditions they split off the carboxyl group. This splitting off of the carboxyl group as carbon dioxide is performed by the so-called carboxylase bacteria, and it may take place either before or after deamidization (see page 615). If it happens after this process, the products are not highly toxic and include phenol, cresol, indole and skatole, which are partly absorbed into the blood and partly excreted with the feces.

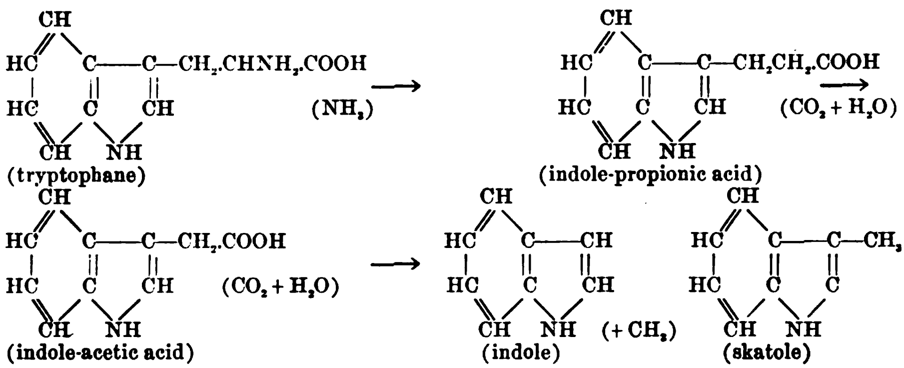
The fractions of those substances that are absorbed into the blood have their toxicity removed by conjugation mainly with sulphuric acid to form the so-called *ethereal sulphates.* A part is also combined with glycuronic acid (see page 632). In the case of phenol and cresol this conjugation occurs immediately after absorption, but in the case of indole and skatole it is preceded by an oxidative process, converting these substances into indoxyl and skatoxyl respectively. The detoxication process occurs in the liver, as has been shown by experiments in which this organ was artificially perfused outside the body. They are then removed from the blood by the kidneys and excreted in the urine. The proportion of ethereal sulphates in this fluid is therefore an indication of the extent of intestinal putrefaction of protein (see page 632). The indican, being readily detectable by the well-known color reaction of Jaffé, serves as an indicator of the extent of intestinal putrefaction. The indole and skatole which are not thus absorbed and detoxicated are excreted with the feces, to which they give the characteristic odor.

The source of the phenol is tyrosine and that of the indole is tryptophane. The chemical processes involved are shown in the following equations, in which the by-products of the reactions are in brackets.



(tyrosine) (p-oxyphenyl-propionic acid) (p-oxyphenyl-acetic acid) (paracresol) (phenol)

Putrefaction of tryptophane is probably preceded by deamidization:



If, however, the carboxylase bacteria remove the carboxyl group before the amino group has been removed, highly toxic substances called amines are produced. They are the so-called *ptomaines*. From alanine, ethylamine is formed; from tyrosine, phenolethylamine; from histidine, which it will be remembered is an important protein building-stone, imidazylethylamine, and so on. The process of formation is illustrated in the accompanying formulæ:

1. $\underset{\text{Alanine}}{CH_3.CH(NH_2).COOH} = CO_2 + \underset{\text{Ethylamine}}{CH_3.CH_2(NH_2)}$
2. $\underset{\text{Tyrosine}}{C_6H_4(OH).CH_2.CH(NH_2).COOH} = CO_2 + \underset{\text{Phenylethylamine}}{C_6H_4(OH).CH_2.CH_2.NH_2}$
3. $\underset{\text{Histidine.}}{C_3N_2H_3.CH_2.CH(NH_2).COOH} = CO_2 + \underset{\text{Imidazylethylamine.}}{C_3H_3N_2.\ CH_2.CH_2.NH_2}$

Similar substances are very common in the metabolic products of plants; for example, they constitute the active principle of ergot. They are also no doubt produced in the tissues of mammals, imidazylethylamine, commonly called histamine, being thus produced, as well as the closely related epinephrine, which is the active principle of the suprarenal gland (see page 737), and may be described as a methylated ethylamine derivative of tyrosine.

Phenylacetic acid produced by a similar process from tyrosine may be excreted in the urine, where it forms the mother substance of homogentisic acid, to which the dark brown color of the urine in alkaptonuria is due.

The great importance attached to these decomposition products of proteins depends on the fact that they have powerful pharmacologic actions. These actions are developed very largely upon the vascular system; histamine, for example, produces marked vasodilatation and lowers the coagulability of the blood, whereas other substances of the

same class, like epinephrine, have the property of raising the blood pressure. In larger doses, serious nervous symptoms and a condition of profound collapse are produced. These observations have led several investigators to believe that the persistent occurrence of bacterial fermentation and the absorption of the resulting decomposition products of protein into the blood ultimately cause arteriosclerosis and the other symptoms that accompany senescence. It is difficult at the present time to know how much of this one ought to believe, although it can not be doubted that putrefaction has an unfavorable action on the arteries, and that an excessive degree of it causes the symptoms of ptomaine poisoning.

If the ptomaines have formed in the food before it is eaten, the symptoms develop in from one to five hours after the meal, but if the decomposition occurs in the intestine on account of bacteria that are taken at the same time as the food, the ptomaines may not have developed sufficiently to cause symptoms until from twelve to forty-eight hours; sometimes, however, they develop in an hour or so. Prominent among the symptoms is usually diarrhea, which develops for the purpose of getting rid of the offending bacteria and ptomaines.

Actual infection of food with bacteria of the paratyphoid-enteritidis type is much more common than poisoning by substances (ptomaines) that have been generated in food before it is taken (Jordan[17]*).* Meat, milk and other protein foods are usually the carriers of the bacilli, and in most of the accurately recorded cases the meat or milk was found to be derived from animals suffering from enteritis or some other infection. Sometimes, however, perfectly good food may become infected by handling. Although the symptoms are usually acute, they may closely simulate those of typhoid fever, and the effects of the attack may linger for weeks or months.

Botulism

The commonest type of poisoning by substances actually present in the food is that known as *botulism.* In this the gastrointestinal symptoms are not pronounced,—indeed, paralysis of the intestinal tract with constipation is the rule,—but those affecting the nervous system, dizziness, diplopia and other visual disturbances, with difficulty in swallowing, are very prominent. The temperature and pulse are usually normal. In practically all of the reported cases of botulism, the source of infection has been food which after having been subjected to some preliminary treatment, such as smoking, pickling, or canning, had been allowed to stand for some time and then eaten without cooking. The Bacillus botulinus, which is responsible for the production of the poisons or toxins, is a

strict anaerobe and is readily destroyed by cooking, as are also the poisons. Antitoxins are formed by sublethal injections. Another but now very rare example of poisoning by products formed in food is that caused by "ergotoxin."

The treatment in such cases is to encourage diarrhea by giving purgatives. If the intoxication is of a more chronic character, the symptoms are vague, consisting of drowsiness, lassitude, headache, and general depression. The treatment here also is to clear out the intestines by a good purge. There can be little doubt that many of the unhealthy conditions of the skin leading to the formation of pimples, acnes, and boils, are also caused by chronic intoxication with protein decomposition products. Again, purgation is the proper treatment.

It is unnecessary in a work of this character to go further into these highly important questions. It is probable, however, that the importance of the relationship of excessive protein putrefaction in the intestine to many of the so-called minor diseases can not be overemphasized. On the other hand, we must be careful not to attribute every sort of chronic condition to this putrefaction. Toxemia is often a shibboleth of the profession. When a chronic disease can not be diagnosed, it is put down as a toxemia. This, however, is not medical science—it is medical shirking. It is certainly unsafe at the present time to conclude that the ordinary symptoms of senescence, such as hard arteries or increased blood pressure, are invariably to be attributed to this cause. It will be remembered that Metchnikoff is largely responsible for such a view, and also that he suggested, as the surest way to ward off the chance of such intoxication, the taking of buttermilk, which would supply bacteria through whose growth in the intestine the protein-destroying bacteria would not be able to thrive. It is probable that the same result could be attained in patients showing undoubted signs of suffering from intestinal putrefaction by a change in diet in the direction of giving more carbohydrate, for, as we have seen, if there is a plentiful supply of this foodstuff in the small intestine, the bacteria do not tend to attack the protein.

Before leaving this subject it is interesting to consider for a moment *the cause of the severe symptoms that follow intestinal obstruction.* This question has recently been diligently investigated by Whipple,[18] who found that the nonprotein nitrogen of blood (page 606) becomes greatly increased in intestinal obstruction. The cause for this increase in nonprotein nitrogen is found to be an excessive breakdown of tissue protein caused by the absorption into the blood of a proteose. When this proteose isolated from obstructed loops of intestine was injected into fasting dogs, profound symptoms of depression were produced, followed, in cases in which the dose was sublethal, by recovery in from twenty-four

to forty-eight hours. Along with these symptoms the nitrogen elimination by the urine increased by 100 per cent. A very interesting fact is that animals can be rendered immune to this proteose by progressively increasing periodic administration. When they are thus immunized, the toxic symptoms do not follow upon its injection, nor are the symptoms produced by artificially creating an intestinal obstruction. Conversely, when a chronic toxic condition is kept up by a partial obstruction, such as that produced by making a gastrojejunal fistula and occluding the duodenum, the animals are less susceptible than normal ones to proteose injection.

We have here and there incidentally referred to *the reaction of various parts of the gastrointestinal contents,* but we would call attention once again to this important subject, especially since many points of uncertainty have recently been cleared up by the accurate observations of Long and Fenger,[19] who used the electrometric method for measuring the hydrogen-ion concentration. The contents of the duodenum removed by means of the Rehfuss tube in man showed a reaction varying from distinctly acid to slightly acid, depending upon the proximity of the tube to the pylorus or papilla, this position being determined by x-ray examination. The slight degree of alkalinity is surprising. Lower down in the duodenum the reaction was as frequently acid as alkaline, the degree of acidity, however, being so slight as to favor rather than retard the digestive powers of the pancreatic juice.

To determine the reaction lower down, the observations were made on recently slaughtered animals (pigs, calves, and lambs), the small intestine being tied off in loops of the upper, middle, and lower thirds. The contents of the last loop were often alkaline, but might be more acid even than those of the first, which were usually faintly of this reaction. Considerable variations were, however, the rule. The mixed intestinal contents of a recently fed dog, removed immediately after death, gave $P_H = 6.79$; i. e., very faintly acid.

DIGESTION REFERENCES

(Monographs)

[1]Pavlov, J. P.: The Working of the Digestive Glands. Trans. by Sir W. H. Thompson, London, Griffin, ed. 2, 1910.
[2]Starling, E. H.: Recent Advances in the Physiology of Digestion, W. T. Keene & Co., Chicago, 1907.
[3]Cannon, W. B.: The Mechanical Factors of Digestion, Internat. Med. Monographs, London, Ed. Arnold, 1911.
[4]Carlson, A. J.: The Control of Hunger in Health and Disease, Univ. of Chicago Press, 1917.
[5]Todd, T. Wingate: The Clinical Anatomy of the Gastrointestinal Tract, Manchester, Univ. Press, 1915.

(Original Papers)

[6]Cannon, W. B., and Cattell, McKeen: Am. Jour. Physiol., 1916, xli, 39.
[6a]Gesell, R.: Proc. Am. Physiol. Soc., Am. Jour. Physiol., 1918, xlv, 559.
[7]Dale, H. H., and P. P. Laidlaw: Proc. Phys. Soc., Jour. Physiol., 1912, xliv, pp. 12, 13.
[7a]Babkin, B. P., Rubaschkin, W. J., and Ssawitsch, W. W.: Arch. f. mikr. Anatomie, 1909, lxxiv, 68.
[8]Macallum, A. B.: Ergeb. der Physiol., xi, 598-657.
[9]Miller, F. R.: Quart. Jour. Exper. Physiol., 1913, vi, 57.
[10]Edkins, J. S.: Jour. Physiol., 1906, xxxiv, 133-144.
[10a]Keeton, R. W., and Koch, F. C.: Am. Jour. Physiol., 1915, xxxvii, 481; also Popielski, L.: Arch. f. d. ges. Physiol., 1901, lxxxvi, 215.
[11]Meltzer, S. J.: Am. Jour. Physiol., 1899, ii, 266.
[12]Cannon, W. B.: Am. Jour. Physiol., 1898, i, 359.
[13]Cannon, W. B., and Blake, J. B.: Am. Surg., 1905, xli, 686. Cf. No. 3.
[14]Macallum, A. B.: See Fitzgerald, M. P., Proc. Roy. Soc., lxxxiii, B, 56.
[15]Harvey, B. C. H., and Bensley, R. R.: Biol. Bull., Wood's Hole, 1912, xxiii, 225.
[16]Long, J. H., et al.: Jour. Am. Chem. Soc., 1917, xxxix, 162 and 1493; also ibid., 1916, xxxviii, 38.
[17]Jordan, E. V.: Food Poisoning, Univ. of Chicago Press, 1917.
[18]Whipple, G. H., Cooke, J. V., and Stearns, T.: Jour. Exper. Med., 1917, xxv, 479. Also Whipple, G. H., Stone and Bernheim: Ibid., 1913, xvii, 286 and 307.
[19]Long, J. H., and Fenger, F.: Jour. Am. Chem. Soc., 1917, xxxix, 1278.
[20]Whipple, C. H., and Hooper, C. W.: Am. Jour. Physiol., 1916, xl, 332 and 349; ibid., 1917, xlii, 257 and 264; Hoope: Ibid., p. 280.
[21]Meltzer, S. J.: Am. Jour. Med. Sc., 1917, cliii, 469.

PART VI

THE EXCRETION OF URINE

CHAPTER LVIII

THE EXCRETION OF URINE

BY R. G. PEARCE, B.A., M.D.

It will be advisable to introduce the subject by a brief review of the essential structural features of the kidney, in so far as they apply to the excretory function of the organ.

STRUCTURE OF THE KIDNEY

The kidney is mainly derived from the surface of the celom, and is a mesodermal structure. In this respect it differs from ordinary secreting glands, which are endodermal in origin. Just as it is more or less unique in its development as a gland, it is also unique in its method of functioning. The physiologic theories of the mechanism of urinary secretion are closely related to the highly characteristic structure of the kidney. For this reason a brief survey of the structure of the different parts of the uriniferous tubules and the epithelial cells with which these are lined, is advisable.

The uriniferous tubule, which is the secreting unit of the kidney, takes its origin in the capsule of Bowman, which may be likened to a hollow sphere of very delicate epithelium, one side of which is invaginated by a very much convoluted capillary mass, the glomerulus. The capsule opens up by a narrow twisted neck into a tubule, which is rather tortuous in the cortex (the proximal convoluted tubule), but soon takes a sharp descending course in the medulla towards the pelvis of the kidney, and doubles back (loop of Henle) in a straight course again to the cortex, where it again makes a twisted course (the distal convoluted tubule), and terminates in a collecting tubule, which, uniting with other tubules, collects the urine and conducts it to the pelvis of the kidney. The capsule is lined with very thin epithelial cells, especially over the capillaries comprising the glomerulus. The proximal and distal tubules

contain epithelium showing a prominent striation. These striations are rows of granules, which run towards the lumen of the cell, becoming less distinct as they approach it and apparently standing in close relationship to the rather prominent internal (lumen) striated border of the cell. Some histologists believe that the striations at the border are

Fig. 170.—Diagram of the uriniferous tubules (*C*) the arteries (*A*), and the veins (*B*) of the kidney.

really cilia, which are described as being immobile. The cilia are shown in Fig. 171. The descending limb of Henle's loop is lined with a thin pavement epithelium with large bulging nuclei. The distal convoluted tubule is lined with cells not unlike those found in the proximal tubules, except that the inner border is not striated. The diameter of the lumen

of the capsule varies with the activity of the kidney, as is shown in the following figures given by Brodie and Mackenzie.[1]

	RESTING KIDNEY MM.	KIDNEY DURING DIURESIS MM.
Mean diameter of capsule	93.4	123.8
" " " glomerulus	90.4	100.0
" " " space of capsule	3.0	23.8
Lumen of proximal convoluted tubule	0.0	17.6
" " distal " "	7.2	20.6

The urinary tubule has a remarkable *blood supply.* The renal arteries arise directly from the abdominal aorta and are very short. They run through the medulla to the cortex, and join with neighboring arteries to

A. *B.*

Fig. 171.—Cross sections of convoluted tubules from kidney of rat. *A*, during slight secretion; *B*, during maximal secretion. (From Sauer.)

form arches from which proceed branches, that radiate into the cortex and give off smaller branches each of which very shortly breaks up into a small capillary tuft,—the glomerulus,—which lies in the invaginated sphere of Bowman's capsule. The capillaries collect into an efferent vessel, which appears to be smaller than the afferent artery, and this vessel in emerging from the capsule again breaks up to form a capillary network about the convoluted tubules, forming their sole blood supply. These capillaries coalesce to form the renal vein. The blood of the kidney must, accordingly, pass through two sets of capillaries.

The kidney is richly supplied with nerves, which are for the most part derived from the celiac ganglion and are in connection with the splanch-

nic and the vagus. Other branches from plexuses in the region of the suprarenal body and the aorta join with those coming from the celiac ganglion to form what is known as the renal plexus, which is arranged in a network along the blood vessels and on the walls of the pelvis of the kidney. These fibers are distributed to the very smallest blood vessels, and nerve fibers have been observed among the cells of the tubules.

THE MECHANISM OF THE EXCRETION OF THE URINE

The great number as well as the variety of substances which are present in both the blood and the urine makes it appear improbable that urine excretion is dependent upon chemical combinations within the renal cells, and leads us to seek a physicochemical mechanism to explain the phenomenon. Can we discover the processes by which the kidney fabricates a highly concentrated solution of salts from a very dilute solution of the same salts in the blood plasma? The problem is complicated by the fact that the ratios existing between the concentration of each urinary salt in the urine and the concentration of the same salt in the blood are different. In other words, the urine is not merely concentrated blood plasma freed from protein.

The passage of water and salts through the capillary wall and through the basement membrane surrounding the renal cell probably takes place by simple diffusion. If it were otherwise, an expenditure of energy would be required, and it is difficult to understand how a basement membrane could bring about energy changes. Any substance to which the cell membrane is permeable will diffuse into the cell until an equilibrium is established between its concentration within the cell and that of the lymph or blood plasma. A nondiffusible substance will not enter the cell because it can not pass through the cell membrane, and if it exerts an osmotic pressure, it will also tend to keep the water in which it is dissolved from entering. If water does pass into the cell under these conditions, it is due to the expenditure of energy opposed to and greater than that which is offered by the osmotic pressure of the nondiffusible substances. Possible sources for such energy are the pressure of the blood in the renal capillaries, which would exert a force opposite to that of its osmotic pressure, and the presence within the cell of a concentration of salts greater than is present in the blood, and able to exercise a sufficient osmotic force to draw fluid into the cell against the osmotic force of the nondiffusible salts. The passage of the urinary constituents through the cell might also be due to simple diffusion, the substances passing through the cell to be extruded on the other side in

the same concentration as in the blood. In this case, the renal cells would act merely as a filter, the urine having the same concentration of each urinary salt as is present in the blood.

A comparison of the concentrations of the urinary salts in the urine and the blood shows, however, that the urine is not merely a deproteinized blood plasma, so that other factors must be sought to explain the excretion. Since the concentration of the urine requires the expenditure of much more energy than is provided by the known physical factors, it is generally accepted that the renal cell in some manner supplies this energy by its metabolic activity. It is impossible at present even to surmise the nature of the process. Two possibilities may be considered. One is that the urine is a filtrate of the blood which has passed through a portion of the renal epithelium into the tubules as a very dilute fluid, resembling the blood plasma minus its colloidal substances, and that this dilute fluid is concentrated by the reabsorption of fluid and of salts by other cells of the kidney, and again replaced in the blood stream. The other is that the salts and fluid are each actively and individually excreted by the kidney. Whichever condition is the true one, the fact remains that the change in the concentration entails the expenditure of a great amount of energy on the part of the renal cells.

The energy which the kidney must use in the actual work of concentrating the urine from the fluid of the blood plasma can not be computed from a comparison of the concentration of the urinary salts as a whole in both the blood and the urine. Each constituent must be considered apart. We can not, for example, determine the molecular concentration of the blood plasma and the urine (by measuring Δ) (page 10) and estimate the work which is expended in producing the concentration from the observed difference. On the basis of such comparisons, however, it is said that the excretion of 100 c.c. of urine requires at the minimum 500 kilogrammeters of work (Cushny[2]). Even this conservative estimate may be wrong, for it does not take into consideration the possibility that the excretion of water by the kidney requires energy expenditure on the part of the renal cells.

Theories of Renal Function

For many years two rival hypotheses have dominated the teaching of the mechanism of renal function. Bowman and Heidenhain postulated that the constituents of the urine are secreted by the vital activity of the epithelium of the capsule and the tubules. The glomerular capsule secretes the water and the easily diffusible salts in a dilute solution, and the uriniferous tubules add to this fluid the various organic and inorganic salts to bring the urine to the necessary concentration. This

theory has been termed the *vital theory*. Ludwig, on the other hand, advanced what is termed the physical theory, which holds that the glomerulus and capsule act simply as a filter, which allows the fluid of the blood plasma to pass through in a very dilute solution and in large amounts. This fluid is concentrated by physicochemical processes on its passage along the urinary tubules to the pelvis of the kidney.

Both of these theories are inadequate and fail to explain the phenomena which research has shown to occur in the kidney, but they have served to develop what Cushny terms *a modern theory* of urinary excretion.

The Modern Theory of Urine Formation.—This theory accepts the general scheme of filtration and reabsorption of Ludwig, but pays due respect to the fact that the known physical forces are not adequate to explain the reabsorption which must occur in the tubules. It therefore supplements Ludwig's theory by assuming a vital activity on the part of the epithelium of the tubules in reabsorbing fluids and salts from the dilute filtrate coming from the glomerulus and capsule. A large amount of plasma fluid is filtered through the walls of the glomerular vessels. This fluid has the same concentration of the salts to which the capsule is permeable as does the blood plasma, but it is free of the colloidal substances normally present in the plasma. The blood leaving the glomerulus is therefore a somewhat concentrated solution of plasma colloids, and must have returned to it the proper amount of water and salts to make it an optimum fluid for the body cells. This is accomplished by active absorption from the glomerular filtrate. The salts that are of no use to the body are not reabsorbed and therefore appear in highly concentrated form in the urine. These salts are termed *nonthreshold substances,* and since their presence in the plasma is unnecessary, they continue to be excreted as long as they are present in any concentration in the blood. The salts that are necessary for the plasma are termed *threshold substances,* and are reabsorbed until they are again present in the plasma in optimal strength. For example, urea continues to be excreted as long as any is present in the blood, while glucose is completely reabsorbed so long as its concentration remains under a more or less fixed level.

It is impossible to give a summary of the arguments which have been advanced in support of any of the theories. However, since the modern theory appears to offer a better explanation of the established facts, it may be wise to recount some of the best experimental evidence in support of it.

First, we must inquire as to the amount of deproteinized blood plasma which the capsule must filter off from the blood in order to furnish the

amount of the various salts excreted each day and the amount of wat
absorbed by the epithelium of the tubules to account for the concentr
tion in which the salts are found in the urine. In order to produce :
grams of urea in 1200 c.c. of urine, 60 liters of blood-plasma fluid co
taining 0.03 per cent of urea would have to be filtered through the ca
sule $\left(\frac{20}{0.03} = 6000\right)$, and 5.9 liters of water returned to the blood fro
the uriniferous tubules. Since the bloodflow through the kidneys is ve
great, at least 500 liters per day, only about 13 per cent of the fluid co
tained in the blood passing through the glomerulus would pass l
filtration through the capsule of Bowman.

The fact that such a large amount of fluid would have to be rea
sorbed from the uriniferous tubules (59 liters) is a possible *a pric*
criticism of the theory, but Cushny points out that the amount ea
tubule would have to absorb per hour would be very small (in his e
periment on a cat amounting to less than 0.014 c.c. per hour).

The filtration of the protein-free blood fluid through the renal capsu
like that through any other membrane, depends on several factors. (
There must be a difference in the pressure between the blood and t
urinary filtrate. In the laboratory the pressure used in filtering
usually supplied by gravity, but in the case of the filtration of the uri
through the capsule the force is furnished by the pressure of blood
the glomerular vessels. (2) The character of the filter determines wh
substances shall pass. The renal capsule is a membrane normally i
pervious to the proteins of the blood, but pervious to the other constit
ents. Under certain conditions it loses this character. (3) The ch
acter of the fluid determines how readily it will filter through the me
brane. If the fluid contains a substance which can not pass through t
filter and which exerts an osmotic pressure in opposition to the filteri
force, the rate of filtration as well as the amount filtered, will be reduc

If the capsule acts as a filter it should be possible to alter the rate
urine excretion by varying any of these factors, and experimentally t
is true. The factors can be varied in several ways. If the blood pressu
is raised by tying off several of the branches of the aorta, the urine
appreciably increased, or if the blood pressure is decreased, as can
done by compressing the renal artery by means of a screw clamp, t
amount of urine is decreased. In the artificially perfused kidney, t
fluid exuding from the ureter increases as the pressure of the perfusi
fluid is increased, and decreases as the pressure is decreased. Whetl
changes in the pressure in the blood are directly responsible for variati
in the rate of urine excretion, or whether they act indirectly by varyi
the rate of the bloodflow in the kidneys, has been the subject of mu

debate. Probably both factors are involved, as is shown by the following observations. If the blood pressure is increased by vasoconstriction in the splanchnic area produced by stimulation of the splanchnic nerves, the flow of blood through the kidney is decreased and the excretion of urine falls. Apparently, secretion can continue only as long as the colloids of the plasma are not notably increased, for, as the osmotic pressure due to the indiffusible colloids rises, the pressure in the capillaries is no longer able to oppose it. The same point has been beautifully shown by Starling and his pupils, who found that the secretion of urine ceases when the capillary pressure in the glomerulus fell below that exerted by the osmotic pressure of the blood proteins, the critical pressure being from 30 to 40 mm. Hg. They also found that dilution of the blood with saline solution by reducing the osmotic pressure of the proteins in the plasma, was accompanied by an increase in the rate of excretion; excretion in such cases being maintained at a blood pressure below the normal critical pressure. If the dilution of the blood was made with saline containing gelatin or gum arabic, on the other hand, the diuretic effect was greatly decreased, and any fall in the blood pressure was followed by a suppression in the urine (Knowlton[9]). These experiments evidently indicate that the saline produces its diuresis by diluting the plasma proteins and lowering their osmotic pressure, since when the osmotic pressure of the blood is maintained by the addition of colloids in which this is present, no diuresis occurs. The significance of these facts, in connection with the raising of lowered blood pressure after hemorrhage, has already been alluded to (page 139).

This view is confirmed by the experiments of Barcroft and Straub,[10] who showed that the oxygen consumption is often not appreciably raised during the diuresis produced by the injection of saline. If the diuresis produced by this means was due to an actual increase in the work of the kidney, the oxygen consumption would have been increased.

In the frog, the glomerulus and the tubules are supplied with blood by the renal artery, as is the case in the mammal, but the tubules curiously enough are also supplied with some of the blood coming from the lower extremities and the trunk through a vessel which has no counterpart in the mammal—the renal portal vein. The blood, therefore, which is supplied to the tubule is a mixture from the glomerulus and the renal portal system. By ligating the renal vessels it is possible to cut off the blood supply of the glomerulus while leaving the tubules supplied by the renal portal vein. Normally the pressure in the renal portal system is not sufficient to force blood back through the glomerular vessels. Ligature of the renal vessels at once results in a suppression of the urine.

If the glomerular vessels are perfused with Ringer's solution at a

pressure equal to that found in the aorta, a considerable flow of fluid may be secured from the ureters, but no fluid is obtained when the renal portal vein is perfused at a pressure equal to that normally present in this vein. Rowntree and Geraghty[11] found that phenolsulphonephthalein added to the perfusion fluid passed through the renal portal vein, did not cause secretion, but when urea was added to the perfusate, fluid containing the dye was obtained from the ureter. Unfortunately the pressure employed in these experiments may have allowed some fluid to be forced backward into the glomerulus, so that the results may be due to filtration through the capsule.

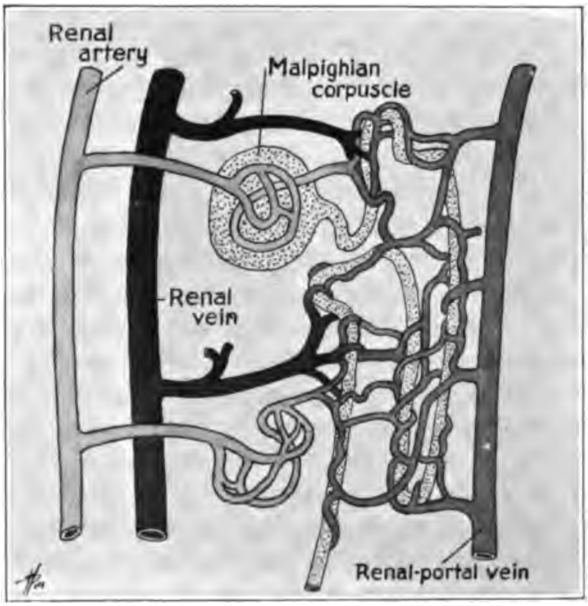


Fig. 172.—Diagram of blood supply of Malpighian corpuscle and of convoluted tubules in amphibian kidney. (Redrawn from Cushny.)

It is generally accepted that the proof that the capsule acts as a filter is fairly complete. Unfortunately such decisive experimental facts can not be offered to prove the assumption that the epithelium of the tubules reabsorbs the excess of water and salts which are filtered off through the capsule. If the modern theory of urine excretion is correct, the cells of the tubules must not only absorb large amounts of water, but they must also allow for the reentrance into the blood, either completely or partially, of certain salts, while they must reject others entirely.

We have called attention above to the fact that the glomerular filtrate is very different from the urine that is finally passed. The urine contains a very high percentage of small molecules, and the proportion in which they

are present is entirely different from that in the blood plasma or in the glomerular filtrate. This is shown in the following figures, which give an average normal value for the urea, uric acid, chlorine, and glucose in 100 c.c. of protein-free blood plasma and 100 c.c. of urine. In the third column is given the change in concentration which has occurred in the kidney.

	100 C.C. PROTEIN-FREE BLOOD PLASMA CONTAINS	100 C.C. URINE CONTAINS	CHANGE IN CONCENTRATION IN THE KIDNEY
Urea	.033	2.	60
Uric Acid	.0022	.05	22.7
Chlorine	.41	.6	1.5
Glucose	.1	—	—

Here the blood plasma fluid contained but 0.033 per cent of urea, and the urine 2 per cent. Accordingly, 6 liters of glomerular filtrate would be required to furnish 100 c.c. of urine, $\left\{\frac{2}{0.33} = 6000\right)$. Six liters of glomerular filtrate would contain 6.6 grams of sugar, 0.132 grams of uric acid, and 24.6 grams of chlorine. But 100 c.c. of urine contains no glucose, 0.05 grams of uric acid and 0.6 grams of chlorine. According to the modern theory, these figures indicate that during the passage of the urine through the tubules 5900 c.c. of water, 6.6 grams of sugar, 24 grams of chlorine and 0.067 grams of uric acid would have to be absorbed by the renal epithelium in the production of 100 c.c. of urine containing the concentration given above.

Among the most convincing experiments that can be offered in support of the absorption of fluid and salts by the tubules, are those in which the pressure of the urine in the tubules is slightly increased by partial closure of the ureter (Cushny). In these experiments the ureter of one kidney is partly closed with a clamp and the excretion obtained from this kidney is compared with that of the opposite normal kidney. In general, obstruction of the ureter results in a decrease in the amounts of water, chloride and urea excreted. But, curiously, the urea content is decreased relatively less than is the chloride and water content. These results can be explained on the basis that any pressure acting to oppose the head of pressure producing filtration in the glomerulus will reduce the amount of the glomerular filtration, and accordingly the time allowed for the passage of this filtrate along the tubules is increased and absorption becomes more complete. Since urea is probably not absorbed at all and chloride is, the discrepancy in the effects on the excretion of urea and chlorine in the partially obstructed kidney can be explained.

When very large amounts of water are taken by mouth, it often hap-

pens that the urine excreted has a concentration of salts less than that present in the fluid of the blood. Some investigators believe that such a condition is possible only on the assumption that water is actively excreted, but a more plausible explanation based on the modern theory is that the water that is absorbed from the alimentary tract reaches the kidney as a dilute saline solution, and is rapidly filtered off in a form somewhat more dilute than the optimal solution which blood plasma must have for the well-being of the tissues. The tubules reabsorb the amounts of water and of those salts, such as chlorides, uric acid, and sugar, necessary to restore the plasma to the optimal concentration, but do not absorb the nonthreshold substances, such as urea.

It is impossible to analyze the forces that are responsible for such a degree of absorption by the epithelium of the tubules. For the present we must classify them, for want of a better term, as vital forces. The questions that await immediate investigation are whether absorption actually takes place, and, if it does so, what factors cause it to vary.

Many attempts have been made, by destroying the capsules or the tubules by means of poisons or by operation, to determine directly or indirectly the question of the function of the tubules.

In such experiments, however, the number of factors involved confuse the issue and make the results practically valueless so far as determining the normal function of the tubules. Other experimenters have attempted to show absorption in the tubules by injecting diffusible substances, such as chemicals and dyes, into the ureter under what they deemed sufficient pressure to force the solution into the tubules, and by an examination of the blood or the tissues to determine whether or not the injected substances had been absorbed. The results obtained by this method are not convincing, probably chiefly because of the difficulty in reaching the tubules. Indeed, it is very questionable whether it is possible to inject a substance into the tubules from the ureter.

Years ago Heidenhain, the exponent of the vital theory of excretion, believed that he had demonstrated the ability of the renal cells to excrete dye substances injected intravenously. Since he failed to find evidence of dye excretion in the capsule, but found masses of dye in the tubules and stained granules in the cells of the tubules, he concluded that the cells of the tubules had the power to excrete the dye, and from analogy he believed that the tubules must likewise excrete the water and the various urinary salts. Subsequent work, however, has failed to confirm his belief that the capsule is not concerned in the excretion of the dye, and it is as reasonable to explain the results of the experiments with the dyes by assuming that the masses of dye substances found in the tubules and in the cells are due to the reabsorption of

water and perhaps of some of the dye from the dilute glomerular filtrate, as to accept Heidenhain's hypothesis.

In the following table taken from Cushny the movements of the constituents of the plasma may be followed through the kidney. The ultimate destination of each is indicated in the enclosures.

	67 LITERS PLASMA CONTAIN		62 LITERS FILTRATE CONTAIN	61 LITERS REABSORBED FLUID CONTAIN		1 LITER URINE CONTAINS	
	PER CENT	TOTAL	IN ALL	PER CENT	TOTAL	PER CENT	TOTAL
Water	92	62 l.	62 l.	—	61 l.	95	950 c.c.
Colloids	8	5360 gm.	——	—	—	—	—
Dextrose	0.1	67 gm.	67 gm.	0.11	67 gm.	—	—
Uric acid	0.002	1.3 "	1.3 "	0.0013	0.8 "	0.05	0.05 gm.
Sodium	0.3	200 "	200 "	0.32	196.5 "	0.35	3.5 "
Potassium	0.02	13.3 "	13.3 "	0.019	11.8 "	0.15	1.5 "
Chloride	0.37	248 "	248 "	0.40	242 "	0.6	6.0 "
Urea	0.03	20 "	20 "			2.0	2.0 "
Sulphate	0.003	1.8 "	1.8 "			0.18	1.8 "

(From Cushny.[2])

It will be noted that the dextrose alone is completely absorbed, and that the urea and the sulphate are not absorbed at all from the glomerular filtrate. The other salts are partly absorbed.

As already mentioned, Barcroft and Straub have shown that the diuresis which results from the injection of saline into the blood is not accompanied by any increase in the oxygen consumption of the kidney. This observation, coupled with the fact that the total amount of chloride, urea, and sulphate which is excreted during saline diuresis, is greater than under normal conditions indicates that the excretion of these salts is not due to any vital secretory power of the kidney, but rather to factors that are extrarenal in origin.

The diuresis produced by adding urea or sodium sulphate to the blood, on the other hand, is accompanied by an increase in the oxygen consumption of the kidney. This increase can not be due to active elimination of these salts by the tubules, the work of which requires oxygen, for no increase in oxygen consumption accompanies the increased excretion of the same salts under saline diuresis. Sulphate and urea are nonthreshold substances, and are not absorbed by the tubules. The explanation of the oxygen consumption is probably that the osmotic pressure which these bodies in the glomerular filtrate exert makes it necessary for the epithelium to oppose a greater absorbing force to concentrate the urine, and hence a greater expenditure of energy is requird.

Diuretics.—The action of the xanthine compounds—caffeine, theobromine and theophylline—in the production of diuresis is unexplained.

It may be due in part to vascular changes and in part to reduction in the resistance to filtration brought about by alteration in the permeability of the capsule.

According to the modern theory the polyuria in diabetes is produced by the excessive amount of water taken and by the inability of the kidney to concentrate the urine against the osmotic pressure offered by the concentrated sugar solution in the tubules. The presence of the hyperglycemia in an amount higher than is present in the optimal blood plasma in this disease makes sugar a nonthreshold substance, so to speak, and none is absorbed. The diuresis following the injection of sugar is therefore of the same type as that produced by sulphate and urea. The diuretic action of the digitalis group is dependent upon its influence on the circulatory system. If the circulation is already sufficient, digitalis does not cause diuresis. The cause of the diuresis produced by pituitary extract is not known. It may be owing in part to its action on the circulation and in part to a direct action on the kidney.

Albuminuria.—The plasma proteins ordinarily do not obtain entrance into the tubules of the kidney. In disease such as acute nephritis and cardiac failure, the plasma colloids are filtered off through the capsule, probably because of some change that has occurred in the permeability of its membrane due to inflammation or asphyxia. In these cases the urine is usually reduced in amount. Probably there is no purely glomerular or tubular type of nephritis, both structures sharing in the disability. While it can not be said that any of the so-called renal tests that have been advanced in recent years are free from criticism, they nevertheless have contributed very useful information. The fact that the kidney of the chronic nephritic excretes a urine of more or less fixed low specific gravity would suggest that here there is an impairment of the resorbing mechanism, and the failure of a kidney to excrete the proper amount of dye, as in the phenolsulphonephthalein test, suggests an impairment in the filtering apparatus. Hard and fast rules can not be applied, however, and probably the tests must at present be interpreted for the kidney as a whole.

The Influence of the Nervous System on the Secretion of Urine.—In spite of numerous and repeated attempts to demonstrate that a nervous mechanism governs the excretion of urine, no proofs which are above criticism have been forthcoming. Stimulation of the splanchnic nerves results in a diminution in the excretion of urine, probably because of a diminution in the blood supply of the renal vessels owing to the vasoconstriction. Stimulation of the vagus nerves below the level of the cardiac branches has been said to result in the augmentation of the rate of urine excretion (Asher and Pearce[12]). The results are doubtful, however, since

there is no increase in the oxygen absorption under the above conditions (Pearce and Carter[13]). In the light of the modern theory this vagal diuresis would be interpreted as due to an inhibition of the absorption in the tubules rather than an augmentation in the actual excretion of urine.

There is no doubt that the renal nerves profoundly affect the excretion of urine, but that they do so directly is very improbable, since perfectly

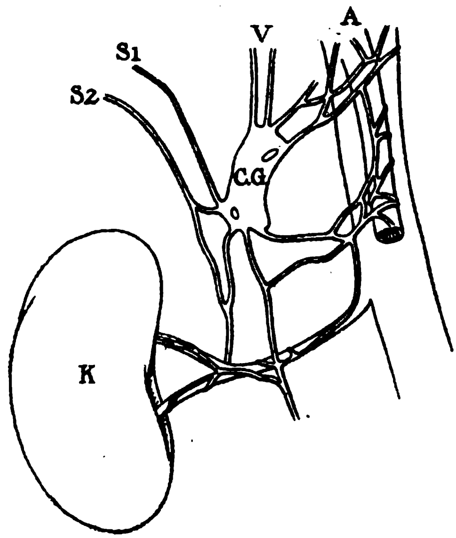


Fig. 173.—Nerve supply of the kidney. *K*, kidney; S_1, S_2, major and minor splanchnic nerves; *V*, vagus; *C.G.*, Celiac ganglion; *A*, aorta. (From Cushny.)

adequate renal function can be maintained in animals that have had the kidneys entirely removed and then replaced. There are numerous reflexes that affect the rate of urine excretion by constriction of the renal vessels. Injury to the bladder or ureter, abdominal injuries to the kidney, or even cold applied to the skin, may result in incomplete suppression of the urine.

CHAPTER LIX

THE AMOUNT, COMPOSITION, AND CHARACTER OF URINE

By R. G. Pearce, B.A., M.D.

In the chapters on digestion and metabolism, we have followed the course which food takes with especial reference to the nutrition of the body. The excretion of these elements of nutrition is taken up under a number of the subdivisions of physiology, viz., respiration, digestion, kidney function and the skin. In the chapters on digestion attention was called to the fact that the feces, besides containing the indigestible residue of the aliment, contain several excretory products which at one time or another have actually been within the body proper. These include normally the pigments of the body and many of the heavier mineral salts, such as iron, magnesium, lime and phosphates; and under abnormal conditions, as when the metals are given as medicine, bismuth and mercury. The respiratory system excretes most of the oxygen and carbon. In this chapter we shall take up the manner in which the body rids itself of the nitrogenous and some of the mineral waste materials. Even at the risk of repetition, it will be advantageous to recapitulate certain facts concerning the essential chemical structure of the urinary constituents, so that we may be in a position to appreciate the kidney function in health and disease.

We now know that the kidney does not form any of the specific constituents of its secretion (except hippuric acid). These substances are formed in the various tissues of the body, and are brought to the kidneys by the blood, where they are eliminated. But while the constituents are unchanged in chemical composition in the urine from that in which they are found in the blood, they do occur in greatly changed proportions. It is this variation in the concentration of the urinary constituents in the blood and the urine which presents the most important and at the same time the most difficult question in the physiology of the kidney. In the following table the percentage composition of the blood plasma is compared with that of an average sample of human urine. The third column gives the change in concentration which each constituent undergoes in passing through the renal filter.

	BLOOD PLASMA PER CENT	URINE PER CENT	CHANGE IN CONCENTRATION
Water	90-93	95	—
Proteins, fats and other colloids	7-9	—	—
Dextrose	0.1	—	—
Urea	0.03	2	60
Uric acid	0.002	0.05	25
Creatinine			
Ammonia	0.001	0.04	40
Sodium	0.32	0.35	1
Potassium	0.02	0.115	7
Calcium	0.008	0.015	2
Magnesium	0.0025	0.006	2
Chlorine	0.009	0.27	30
Phosphates (PO_4)	0.003	0.18	60
Sulphates (SO_4)			
Amino acids			

The Amount of Urine

The amount of urine passed in twenty-four hours varies with the amount of fluid ingested and the proportion of fluid retained by the body or excreted by other channels. Under ordinary conditions a twenty-four-hour sample amounts to from 1000 to 1800 c.c. of urine. On a constant water intake the volume of urine is extremely variable for any single day or part of the day (Addis and Watanabe[3]). The average volume of urine excreted by twenty individuals on the third, fourth and fifth days of a constant diet in which the fluid intake was 2,070 c.c., varied from 1,013 to 1,712 c.c. for a twenty-four-hour period, from 684 to 1,195 c.c. for the first twelve hours of the day, and from 501 to 788 c.c. for the first eight hours of the day. In normal subjects the amount of urine excreted during the night is usually less than that during the day. This is such a constant finding that in cases where more than 50 per cent of the urine is excreted in the twelve hours of the night, suspicions of renal disease should be aroused.

The Specific Gravity of Urine

In urine collected at different times of the day the specific gravity may show a variation of ten points. Indeed, the specific gravity of the urine has been taken as a functional test by clinicians. With a constant food and water intake the variations found in the specific gravity of samples of urine taken at two-hour periods in normal and pathologic conditions are very useful as criteria of the functional state of the kidney. Fixation of the specific gravity at either a low or a high figure is not the usual normal finding. The following figures will illustrate:

	DAY						NIGHT
	8-10 A.M.	10.12 A.M.	12-2 P.M.	2-4 P.M.	4-6 P.M.	6-8 P.M.	8-8 P.M.-A.M.
Normal person	1.016	1.019	1.012	1.014	1.020	1.010	1.020
In Hypertensive Nephritis	1.010	1.009	1.010	1.009	1.019	1.010	1.009
In Myocardial Decompensation	1.018	1.020	1.019	1.018	0.020	1.021	1.022

(Compiled from Mosenthal's figures.)

The proportion of water to total solids is often very similar in plasma and urine, but when water is taken in large quantities the urine shows much greater changes than does the blood, and the solids may sink to a very low concentration. On the other hand, when little fluid is taken or when the skin and bowel eliminate a large amount of fluid, the urine may become very concentrated without any change in the blood plasma. The total solids in urine can be determined with approximate accuracy by multiplying the last two figures of the specific gravity by the constant coefficient 0.233 (Haeser).

The Depression of Freezing Point

While the solids of the blood consist, for the most part, of proteins and colloids, those of the urine are made up of inorganic salts and small organic molecules. The molecular concentration—that is, the total number of molecules in a given quantity of fluid—is under ordinary conditions much greater in the urine than in the blood. The molecular concentration may be determined by the depression of the freezing point of a fluid below that of distilled water (see page 10). Blood freezes almost constantly at –0.56° C., while urine may freeze at variations of temperature between –1° C. and –2.5° C.; if very concentrated it may freeze at a temperature as low as –5° C., or if dilute the freezing point may be as high as –0.075° C.

The variability of the freezing point and the specific gravity of the urine lead us to a consideration of the relationship of the urinary volume to its concentration. In the first place, the volume of water ingested is more frequently than otherwise in excess of the minimum absolutely required by the body, and is subject to greater variation than the substances excreted in the urine. The kidney is able to eliminate one constituent of the plasma which may be present in excess without involving any changes in others. For example, when salt is added to the food and excreted in the urine, the total chlorides are increased, but the amount of urine and the other constituents may remain unchanged; or, again, as may happen, excess of salt leads to an increase in the volume of the urine, but the salt concentration remains constant while that of the other urinary bodies is decreased. Similarly, although the rate of urea

excretion is not demonstrably augmented by an increase in the volume of the urine, an increase in the rate of urea excretion induced by the ingestion of urea is accompanied by a larger volume of urine. That these two factors may not stand in a causal relationship to each other is suggested by recent work of Addis and Watanabe,[3] who find no quantitative relationship between the rate of increase in urea excretion and the increase in urine volume, and who believe that the apparent relationship is due to a common cause, such as alteration in the rate of circulation or change in the activity of the kidney cells. Nevertheless, there appears to be a limit set to the power of the kidney to take the urinary salts or water from the plasma and to place them in the urine in quite different proportions. The definite amount of water required to hold the urinary salts has been termed the "volume obligative" (Ambard[5]). These limits of concentration may be fixed by the energy which the kidney can bring to act against the osmotic resistance.

The inconstancy in the behavior of the kidney toward ingested salts is probably due to the fact that the salts reach the kidney in the concentration in which they are held by the blood plasma, and not as they were ingested. If salt is absorbed rapidly enough to disturb the salt equilibrium of the tissues and plasma, then water will be abstracted from the tissues, and the plasma on reaching the kidney will eliminate the salt and water together. The difference in the reaction arises from the varied activity in the tissues in general rather than in the kidney itself.

The Reaction of Urine

In man and the carnivora this reaction is generally acid to litmus or phenolphthalein. The cause is found in the fact that the end products of protein metabolism give rise to sulphuric and phosphoric acids the acidity of which gives the urine an acid reaction. In the herbivorous animals the alkaline reaction is due to the fact that vegetables and fruits contain salts of dibasic or polybasic acids, such as acid potassium malate, citrate, acetate, and tartrate. Oxidation of these in the body gives rise to carbonates. Some of the carbonic acid is excreted through the lungs, and hence the associated base, generally sodium or potassium, is combined so as to form a weak basic salt.

The measurement of the acidity of the urine in terms of gram anions or cations, like the same measurement in blood, requires the use of the rather difficult electrical or indicator method, the principle of which has been described in Chapter V. Expressed in terms of C_H, the acidity varies between 4.7×10^{-7} and 100×10^{-7}. *The total potential acidity*—that is, the number of H ions which will be formed in the face of a con-

tinual neutralization of those in solution—may be obtained fairly accurately by titrating the urine with 1/10 normal alkali in the presence of neutral potassium oxalate, using phenolphthalein as an indicator (Folin). The results may be expressed in acidity per cent in terms of c.c. N/10 NaOH required to neutralize 100 c.c. of urine. If the ammonia excretion is added to the titration results, the total potential acidity is very closely measured.

The urine is more alkaline shortly after meals than at other times, since acid is being excreted by the gastric glands. It is more acid on a meat than on a vegetable diet, and is acid during starvation because protein is then the chief metabolite. In disease there is no characteristic variation, save that the urine is more generally acid, which may be explained by the fact that in serious illness the diet is restricted. When the acidity is increased, the excretion of ammonia is usually greater, since ammonium carbonate, the forerunner of urea, acts as an alkali and neutralizes the acid radicles. This rise in ammonia, however, is not always proportional to the acid radicles present, since the fixed alkali derived from fruits and vegetables may be sufficient to neutralize the acid formed.

THE SOLID CONSTITUENTS

For practical reasons we shall divide the constituents of the urine into normal and abnormal. The former are present in the average urine in amounts sufficient to be detected by ordinary means; the latter only rarely appear in detectable quantities. In a person eating an ordinary diet the most important organic and inorganic constituents of the urine are as follows:

TOTAL SOLIDS (40 TO 60 GRAMS) IN ONE LITER OF NORMAL URINE

ORGANIC CONSTITUENTS, 25-40 GM.	INORGANIC CONSTITUENTS, 15-25 GM.
Urea, 20-35 gm.	Sodium chloride ($NaCl$), 8-15 gm.
Creatinine, 1.011.5 gm.	Phosphoric acid (P_2O_5), 2.5-3.5 gm.
Uric acid, 0.5-1.25 gm.	Sulphuric acid, (SO_3), 2-2.5 gm.
Hippuric acid, 0.1-1.7 gm.	Potassium (K_2O), 2-3 gm.
Other constituents (ethereal sulphates, oxalic acid, urinary pigments, etc.), 1.5-2.3 gm.	Sodium (Na_2O), 4-6 gm.
	Calcium (CaO), 0.1-0.3 gm.
	Magnesium (MgO), 0.2-0.5.
	Ammonia (NH_3), 0.3-1.2 gm.
	Iron (in pigment), 0.001-0.010.

(Compiled from Mosenthal's[4] figures.)

These urinary salts are present in the blood, and are excreted only by the kidney. An investigation of the mechanism of renal secretion must therefore include a study of the relationship existing between the concentration of the urinary salts in the blood and in the urine.

The Normal Organic Salts of the Urine

Nitrogenous Constituents.—The greater number of the organic salts of the urine are made up of bodies which contain nitrogen, and which are derived from the protein element of nutrition. The proteins, which form the chief building material of the body, are broken up into their constituent amino acids in the intestinal tract and absorbed as such by the blood. Portions of these acids are taken up by the tissues to repair and to replace those proteins which have been discarded, and the remaining protein, in excess of the body need for amino acids, is deamidized, the major portion of the carbon, oxygen and hydrogen being oxidized to form CO_2 and water, and the lesser portion of these elements being combined with the nitrogen to form urea, ammonia, uric acid, etc. A similar fate later awaits the nitrogen moiety which found a place in the tissues, and which is replaced in turn by new nitrogenous bodies.*

Since all the ingested nitrogen, except a small and rather constant amount which is lost by the feces and the sweat, is excreted in the urine, the total nitrogen of the urine has been taken as a measure of the nitrogen or protein metabolism of the body. In normal conditions the protein metabolism is adjusted in such a manner that the nitrogen intake is equal to the nitrogen output, a condition known as nitrogenous equilibrium. If the nitrogen intake is reduced below the actual body needs, the excretion of nitrogen is greater than the intake which indicates that the body protein is replacing the protein usually furnished by the food. The minimum amount of protein that the body must have to maintain equilibrium varies in individuals, but is on the average between 5 and 6 grams of nitrogen a day, which corresponds to about 40 grams of protein. With the ordinary diet it is usually between 12 and 20 grams a day, or represents from 75 to 125 grams of protein. Since protein is not stored by the body except in periods of growth or after periods of undernutrition, an increase in the protein food is accompanied by an increase in the nitrogen excreted in the urine. For this reason, unless the amount of nitrogen ingested is known, the study of the total nitrogen of the urine gives no information concerning the nature of the nitrogen metabolism of the body. The total output of nitrogen per day usually amounts to 10 to 15 grams—from 1 to 2 per cent of the urine by weight.

All the nitrogenous bodies of the urine are normally nonprotein, and arise from similar bodies in the blood, where they exist in concentrations of from 20 to 30 mg. per 100 c.c. In excreting the nitrogen of the urine the kidney therefore takes it from a solution in which it is found in a concentration of 0.03 per cent on the average and delivers it to a

*For further details see page 610.

solution containing an average of 1.00, or concentrates it at least times.

Urea.—The chief of the nitrogenous bodies of the urine is urea, t origin of which has been fully described in the chapters on metabolis No constituent of the urine is subject to greater variation both in ab: lute and in relative amounts. On an average diet containing 120 grai of protein per day, the absolute urea excretion may amount to about grams; on a low protein diet it may be only a few grams. When the pı tein intake is high, the nitrogen eliminated as urea may be 90 per ce of the total nitrogen; but when the protein intake is low, this proporti may fall to 60 per cent. The difference is because on a low protein d: the greater percentage of nitrogen eliminated is endogenous in orig: and urea, which is the chief constituent of the exogenous nitrogen moie of the urine, is accordingly decreased on low diets.

In recent years the importance of the relationship between the cc centration of the urinary constituents in the blood and the urine h been much insisted upon, and since the estimation of the amount urea in the blood and the urine is relatively simple, most of the wo has been done by using these values. Ambard and Weil[5] believe that quantitative relationship exists between the rate of urine excretion a: the concentration of urea in the blood and the urine, since the urea the blood acts as a stimulus to the renal cells. By comparing the ra of urea excretion and the concentration of urea in the blood and uri in a mathematical formula, they have obtained a value which they l lieve is more or less fixed for the normal kidney. This expression known as *Ambard's coefficient and formula,** and has been used as means of evaluating the functional capacity of the kidney.

Whatever the value of the formula may be in expressing the relationsh existing between the rate of urea excretion and the concentration of tl salt in the blood, it is certain that, in diseased conditions where impa ment of the kidney is certain, the concentration of urea in the blood ı mains permanently at an abnormally high average level, although t

*Ambard and Weil's formula is:

$$K = \frac{Ur}{D \times \frac{70}{P} \times \frac{\sqrt{C}}{\sqrt{25}}}, \text{ in which:}$$

K = coefficient of urea excretion (Constant of Ambard).
Ur = grams of urea per liter of blood.
D = output of urea in grams per 24 hours.
P = weight of the patient.
C = grams of urea per liter of urine.
70 = standard weight.
25 = standard concentration of the urine.

The average value for this constant in normal individuals is said to lie between .06 and .09. Critical reviews of the work have been published recently by Maclean[6] and by Addis : Watanabe.[8]

amount of urea excreted during twenty-four hours may be exactly the same as under normal conditions. Probably the increased concentration of urea in the blood under these conditions is a compensatory measure to provide sufficient pressure to cause its excretion through a damaged outlet. It is this increase in urea of the blood which is indicated by the term *urea retention* in nephritis.

It must not be lost sight of, however, that the approximate constancy of the combined formula is due in large part to the mathematical construction, and also to the fact that any increase in the concentration of urea in the blood is usually accompanied by an increased rate of urea excretion. The factors which are most variable occur as the square or the square roots of their values, and thus the disturbing effect they produce on the constancy of the resultant of the formula is greatly reduced, while the most constant factor, the concentration of urea in the blood, is used with modification. In such a complex mechanism as the renal function it is very probable that other factors are of great importance in controlling the rate of urinary excretion. Many of these factors can not admit of mathematical expression. The writer seriously doubts the advisability of adopting an empirical formula as a means of expressing unknown physiologic laws. Such measures are apt to give a sense of knowledge altogether false, and thus hinder research progress.

The upper limit of blood urea-nitrogen is about 20 mg. per 100 c.c., which would correspond to about 0.45 gm. of urea per liter of blood. The average figure is half of this amount. The maximum concentration of urea in the urine is seldom over 8 per cent. On this basis the kidney can raise the concentration of the urea in the urine, at a conservative estimate, from 100 to 200 times. Normally the daily output of urea nitrogen may range from 8 to 12 gm., and the nitrogen which it contains is roughly 80 per cent of the total excretion for the day.

Ammonia.—The chief source of ammonia in the body is from the nitrogenous portion of the deamidized amino acids. The ammonia found in excess in the portal blood is derived from ingested ammonium salts and from ammonia resulting from bacterial action on proteins in the intestinal tract. The ammonia of the body is present chiefly in the form of ammonium carbonate, and it is this salt that is the precursor of urea. Because ammonium carbonate is so readily converted into urea by the tissues of the body, little ammonia is normally present in the systemic blood. The greater portion of the ammonia that finds its way into the urine serves as a base to transfer acid radicles either ingested or formed within the body. The amount of ammonia in the urine, therefore, is an indirect measure of the extent of urea formation and of the acid bodies

of the blood. For the latter reason the determination of the ammonia excretion in urine is of some clinical importance. The ingestion of mineral acids increases the ammonia excretion, while alkalies tend to reduce it. During fasting and in diseases such as diabetes, where there is an abnormal metabolism, the amount of ammonia in the urine is increased. Ordinarily the daily output of ammonia nitrogen does not exceed 0.5-0.6 gm., constituting 3-5 per cent of the total amount of nitrogen.

Creatinine.—On a meat-free diet the daily excretion of creatinine is remarkably constant, amounting to from 7 to 11 mg. per kilogram of body weight. For this reason its determination is accepted as an indispensable feature in metabolism investigations involving urine analysis.

Any gross variation from the normal amount indicates the certain failure of the attendants to collect all of the twenty-four-hour specimen of urine. Normally the blood contains from 1 to 2 mg. per 100 c.c.

The creatinine is one of the last of the urinary constituents to accumulate in the blood during renal insufficiency, and for this reason affords a reliable prognostic indication concerning the patients' condition. A rise in the creatinine concentration of the blood is evidence of serious renal disease, patients with concentrations of 5 mg. never recovering (Chase and Meyers)[7] The concentration of creatinine in the urine is about 100 times greater than in the blood.

In adult man creatine does not appear in the urine save during starvation or wasting diseases. In woman it is absent save after postpartum resolution of the uterus. Children commonly excrete creatine along with creatinine until the middle years of childhood.

The Purine Bodies and Uric Acid.—The most important purine in human urine is uric acid. Xanthine is the next in importance, and small amounts of hypoxanthine, guanine, and adenine are found. Among the most interesting of the salts of the urine to the clinician are the urates, because an accumulation of uric acid in the body was believed to be responsible for many obscure clinical conditions. It is quite true that the salts of uric acid are found in higher than normal amount in some diseases, especially gout, leukemia, and chronic nephritis, but the many vague theories associated with uric acid and disease have long ago been exploded.

The human body has the almost unique distinction among mammals of not being able to destroy any of the uric acid it produces, and hence all the uric acid formed during metabolism must be excreted in the urine. Unfortunately the kidney appears to be less competent to rid the body of this waste than it is of the other urinary metabolites, and one of the

earliest signs of renal insufficiency is now held to be a failure of the kidney to prevent the uric acid of the blood from increasing. Perhaps the reason for the inability of the kidney to excrete uric acid readily lies in the fact that its salts are among the least soluble of those in the urine. It is on this account that when the urine cools, a red sediment of urates containing certain pigments often separates out.

The uric acid of the urine is possibly derived entirely from the purine metabolism of the body, in which the nucleins either of the body cells or of the exogenous food take part. It is decreased during starvation and increased by eating food rich in nucleins, such as liver and sweetbreads.

Under ordinary conditions the excretion of uric acid amounts to from 0.3 to 1.2 gm. per day (0.02 to 0.10 per cent), the variation being dependent upon the state of health, diet, or personal idiosyncrasy. The blood of a normal individual contains on the average 1.8 mg. of uric acid per 100 c.c. The kidneys are therefore able to concentrate the uric acid in the urine from 30 to 60 times over its concentration in the blood plasma.

The purines found in coffee and tea (caffeine, etc.) are excreted in the urine as salts not of uric acid but of methylated xanthines.

Hippuric Acid.—This is a constant constituent of the urine of herbivorous animals, and is usually present in small amounts in human urine. The amount rarely exceeds 0.7 gm. a day, but on a diet rich in fruits and vegetables it may exceed 2 gm. It is interesting, since it is the only urinary constituent that is synthesized by the renal cells.

Amino acids are always present in small amounts in the urine, constituting, according to D. D. Van Slyke, about 1.5 per cent of the total nitrogen. The estimation of the amino-acid nitrogen of the urine has not been found to be of any clinical significance.[8]

The aromatic oxyacids are normally present in the urine in varying amounts. These include phenol, indoxyl, skatoxyl, and phenylacetic, paraoxyphenyl, propionic, oxymandelic and homogentisic acids. These bodies are derived from phenylamino acids, such as tyrosine, tryptophane, and phenylalanine. It is believed that the putrefactive decomposition of proteins in the large intestine results in the production of these toxic bodies. The body protects itself by oxidizing them and uniting them to sulphuric acid to form the ethereal or conjugated sulphates, which are found in the urine in the form of sodium or potassium salts. The determination of the amounts of these bodies in the urine has therefore been taken as an index of the putrefaction going on within the bowel.

The chief of these bodies is urinary *indican,* which is found usually as a potassium salt. The test for indican in the urine consists in oxidiz-

ing the indoxyl in an acid solution by means of ferric chloride to indigo blue, and shaking out the indigo blue with chloroform. The depth of the color of the chloroform affords a rough means of determining the amount of indican present. The fact that the indican test is negative must not be taken to mean that the intestinal processes are normal, for if the intestine fails to contain phenylated amino acids, or the proper bacteria are not present, no indican will be found. On the other hand, the putrefactive process of the large bowel may not be very extensive, yet the amount of indican in the urine be increased, because of greater absorption due to constipation.

Skatole, a fecal-smelling substance, is formed by certain kinds of bacteria. The greater proportion of this substance is excreted by the bowel, but if the person is constipated, some of it may find its way into the blood to impart a fecal odor to the breath and urine. Its presence therefore has some diagnostic importance.

A very interesting body which is sometimes found in the urine is *homogentisic acid.* It is thought to be an intermediate step in the metabolism of tyrosine, and is found in the urine of people suffering from alkaptonuria. The disease is remarkable in that it appears to run in families and produces no ill effects. Homogentisic acid is a strong reducing agent, and for this reason may be confused with sugar in Fehling's test.

The inorganic constituents of the urine include the acids: chlorides, sulphates and phosphates; and the bases: sodium, potassium, magnesium, and calcium.

The Acids of the Urine.—The *chlorides* compose the bulk of the acid radicles in the urine. Although they appear to be necessary constituents of the living cell, they do not, so far as known, enter into combinations with the organic constituents. The tissues appear to require a rather definite concentration of sodium chloride in order to carry on their work, for reduction in the sodium-chloride intake of the body results in a reduction in the chloride excretion by the urine. In salt starvation the chlorides may disappear entirely from the urine, the amount of chloride excreted appearing to be closely related to the amount of salt ingested. When the intake is constant, the rate of excretion is likewise more or less constant, but a sudden reduction in the salt of the diet may be accompanied by a slight decrease in the salt content of the blood, with an attendant loss of water. On the other hand, when the salt is again taken, there is a retention of salt and of water, with a consequent increase in body weight, until equilibrium is re-established on the old level. While the above is the usual reaction, a considerable retention of salt without an increase in the water content of the body may occur in

some apparently normal cases. This is due probably to the deposition of salt in the tissues.

Careful studies fail to confirm the idea that there is a fixed relationship between the salt and the water of the body. As with the nitrogenous constituents, however, there appears to be a relationship between the rate of excretion of chlorides and the amount of chloride in the blood. Ambard believes that this relationship, like that of the excretion of urea to the blood urea, is capable of being expressed mathematically (see page 527), if allowance is made for the fact that NaCl is not excreted after it falls below a certain concentration in the blood equal to about 5.62 gm. per 1000 c.c. This level is more or less constant for normal individuals, but is considerably increased in disease of the kidney. This is known as the threshold of chloride excretion.

The amount of sodium chloride excreted in the urine in twenty-four hours varies between 8 and 20 gm. a day, according to the intake. It is therefore apparent that the kidney is able to concentrate the salts of the plasma from ten to twenty times.

The Sulphates.—Since the inorganic sulphates do not form an important constituent of the food, the greater portion of the sulphates of the urine are derived from the sulphur found in the protein molecule. For this reason the sulphates of the urine, like the nitrogen, are a measure of protein metabolism. An increase in the nitrogen excretion is accompanied by an increase in the sulphur excretion, the ratio being about 5 to 1. The daily output of sulphur is between 1 and 3 gm. The greatest output is in the form of the alkaline sulphates, about 10 per cent in combination with aromatic bodies, and a small amount in combination with amino acids and neutral organic salts.

The *phosphates* of the urine are derived from the food and from the oxidation of phosphorus-containing bodies in the tissues such as nuclein, lecithin, etc. The daily excretion varies between 1 and 5 gm., calculated as P_2O_5. When calcium or magnesium is present in the food, they are excreted by the bowel as phosphate, and proportionately less is found in the urine. The amount usually excreted in the feces equals about 30 per cent of the total.

Since phosphates in the urine exist as a mixture of the mono- and disodium hydrogen phosphates, they have an important bearing on the reaction of the urine, the amount of each varying with the degree of the acidity of the urine.

On a heavy protein diet the urine is acid on account of the sulphuric and other acids formed from the meat, and in this case there is a greater amount of phosphoric acid and the mono-sodium hydrogen phosphate. When the urine is alkaline or less acid, as it is on a vegetable diet, there

is a large amount of the disodium hydrogen phosphate. Since calcium and magnesium phosphates are more soluble than the diphosphates of the same metals, deposits of the earthy phosphates are often found in neutral or alkaline urines. When the urine is heated, the diphosphate of calcium breaks up into the mono-calcium and a tri-calcium phosphate, which accounts for the fine turbidity often taken for albumin in the flame test. Addition of acid will cause this to disappear. The crystals of triple phosphates which occur in alkaline urine are ammonium magnesium phosphate, NH_4MgPO_4.

KIDNEY REFERENCES

(Monographs)

Beddard, A. P.: Recent Advances in Physiology, Longmans, Green & Co., London, 1906.

Cushny, A. R.: Secretion of Urine, Longmans, Green & Co., London, 1917.

(Original Papers)

[1]Brodie, T. G., and Mackenzie, J. J.: Proc. Roy. Soc., 1914, lxxxvii, B, 593.

[2]Cushny, A. R.: Secretion of Urine, 1917, p. 48.

[3]Addis and Watanabe: Jour. Biol. Chem., 1916, xxiv, 203.

[4]Mosenthal, H. O.: Arch. Int. Med., 1915, xvi, 733.

[5]Ambard and Weil: Physiologie normale et pathologique des reins, Paris, 1914, J. B. Bailliere et fils.

[6]Maclean, F. C.: Jour. Exper. Med., 1915, xxii, 212.

[7]Chase and Meyers: Jour. Am. Med. Assn., 1916, lxvii, 931.

[8]Van Slyke, D. D., and Meyer, G. M.: Jour. Biol. Chem., 1912, xii, 399; and 1913, xvi, 197, 213 and 231.

[9]Knowlton, F. P.: Jour. Physiol., 1911, xliii, 219.

[10]Barcroft, J., and Straub, H.: Jour. Physiol., 1910, xli, 145.

[11]Rowntree and Geraghty: Jour. Pharm. and Exper. Therap., 1910, i, 579.

[12]Asher and Pearce, R. G.: Zeitschr. f. Biol., 1913, lxiii, 83.

[13]Pearce, R. G., and Carter, E. P.: Am. Jour. Physiol., 1915, xxxviii, 350.

PART VII

METABOLISM

CHAPTER LX

METABOLISM

Introductory.—The object of digestion, as we have seen, is to render the food capable of absorption into the circulatory fluids—the blood and lymph. The absorbed food products are then transported to the various organs and tissues of the body, where they may be either used at once or stored away against future requirements. After being used, certain substances are produced from the foods as waste products, and these pass back into the blood to be carried to the organs of excretion, by which they are expelled from the body. By comparison of the amount of these excretory products with that of the constituents of food, we can tell how much of the latter has been retained in the body, or lost from it. This constitutes the subject of *general metabolism*. On the other hand, we may direct our attention, not to the balance between intake and output, but to the chemical changes through which each of the foodstuffs must pass between absorption and excretion. This is the subject of *special metabolism*. In the one case we content ourselves with a comparison of the raw material acquired and the finished product produced by the animal factory; in the other we seek to learn something of the particular changes to which each crude product is subjected before it can be used for the purpose of driving the machinery of life or of repairing the worn-out parts of the body.

In drawing up a balance sheet of general metabolism, we must select for comparison substances that are common to both intake and output. In general the intake comprises, besides oxygen, the proteins, fats and carbohydrates; and the output, carbon dioxide, water and the various nitrogenous constituents of urine. This dissimilarity in chemical structure between the substances ingested and those excreted limits us, in balancing the one against the other, to a comparison of the smallest fragments into which each can be broken by chemical agencies. These are the elements, and of them carbon and nitrogen are the only ones which it is possible to measure

with accuracy in both intake and output. From balance sheets of intake and output of carbon and nitrogen and from information obtained by observing the ratio between the amounts of oxygen consumed by the animal and of carbonic acid excreted, we can draw far-reaching conclusions regarding the relative amounts of protein, fat and carbohydrate that have been involved in the metabolism.

As has already been stated, the essential nature of the metabolic process in animals is one of oxidation—that is, one by which large unstable molecules are broken down to those that are simple and stable. During this process of *catabolism,* as it is called, the potential energy locked away in the large molecules becomes liberated as actual or kinetic energy—that is, as movement and heat. It therefore becomes of importance to compare the actual energy which an animal expends in a given time with the energy which has meanwhile been rendered available by metabolism. We shall first of all consider this so-called *energy balance* and then proceed to examine somewhat more in detail the *material balance of the body.*

ENERGY BALANCE

The unit of energy is the large calorie (written C.), which is the amount of heat required to raise the temperature of one kilogram of water through one degree (Centigrade) of temperature.* We can determine the calorie value by allowing a measured quantity of a substance to burn in compressed oxygen in a steel bomb placed in a known volume of water at a certain temperature. Whenever combustion is completed, we find out through how many degrees the temperature of the water has become raised and multiply this by the volume of water in liters. Measured in such a *calorimeter,* as this apparatus is called, it has been found that the number of calories liberated by burning one gram of each of the proximate principles of food is as follows:

Carbohydrates	Starch	4.1
	Sugar	4.0
Protein		5.0
Fat		9.3

The same number of calories will be liberated at whatever rate the combustion proceeds, provided it results in the same end products. When a substance, such as sugar or fat, is burned in the presence of oxygen, it yields carbon dioxide and water, which are also the end products of the metabolism of these foodstuffs in the animal body; therefore, when a gram of sugar or fat is quickly burned in a calorimeter, it releases the same

*The distinction between a calorie and a degree of temperature must be clearly understood. The former expresses *quantity* of actual heat energy; the latter merely tells us the intensity at which the heat energy is being given out.

amount of energy as when it is slowly oxidized in the animal body. But the case is different for proteins, because these yield less completely oxidized end products in the animal body than they yield when burned in oxygen; so that, to ascertain the physiologic energy value of protein, we must deduct from its physical heat value the physical heat value of the incompletely oxidized end products of its metabolism. It is obvious that we can compute the total available energy of our diet by multiplying the quantity of each foodstuff by its calorie value.

Methods.—In order to measure the energy that is actually liberated

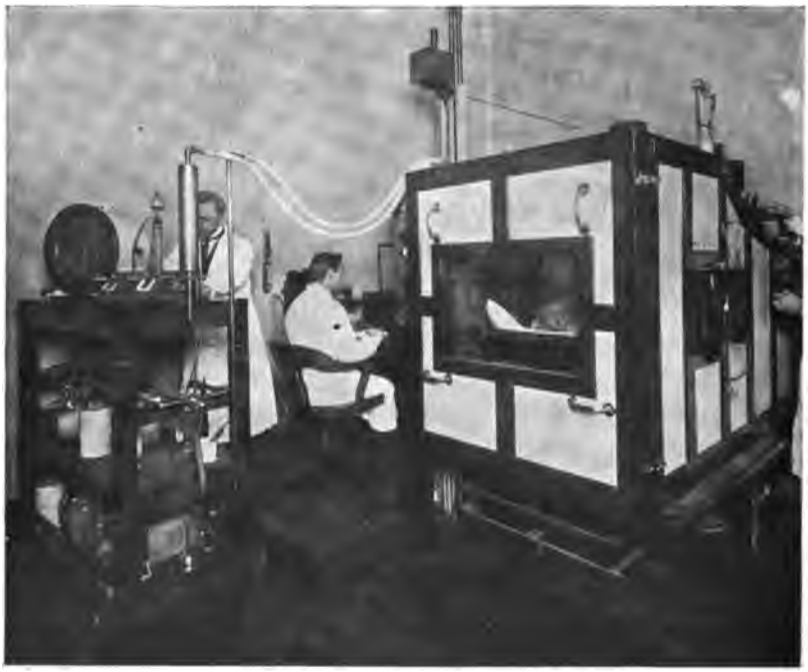


Fig. 174.—Respiration calorimeter of the Russell Sage Institute of Pathology, Bellevue Hospital, New York. At the right is seen the table with the absorption tubes; and in the middle, at the back, the electric control table for regulating the temperature of the double walls of the calorimeter. At the extreme left is the oxygen cylinder. (Lusk's *Science of Nutrition.*)

in the animal body, we must also use a calorimeter, but of somewhat different construction from that used by the chemist, for we have to provide for long-continued observations and for an uninterrupted supply of oxygen to the animal. *Animal calorimeters* are also usually provided with means for the measurement of the amounts of carbon dioxide (and water) discharged and of oxygen absorbed by the animal during the observation. Such respiration calorimeters have been made for all sorts of animals, the most perfect for use on man having been constructed in America (see Fig. 174). As illustrating the extreme accuracy of even the largest of these,

it is interesting to note that the actual heat given out when a definite amount of alcohol or ether is burned in one of them exactly corresponds to the amount as measured by the smaller bomb-calorimeter. All of the energy liberated in the body does not, however, take the form of heat. A variable amount appears as mechanical work, so that to measure in calories all of the energy that an animal expends, one must add to the actual calories given out, the calorie equivalent of the muscular work which has been performed by the animal during the period of observation. This can be measured by means of an ergometer, a calorie corresponding to 425 kilogram* meters of work. That it has been possible to strike an accurate balance between the intake and the output of energy of the animal body, is one of the achievements of modern experimental biology. It can be done in the case of the human animal; thus, a man doing work on a bicycle ergometer in the Benedict calorimeter gave out as actual heat 4,833 C., and did work equalling 602 C., giving a total of 5,435 C. By drawing up a balance sheet of his intake and output of food material during this period, it was found that the man had consumed an amount capable of yielding 5,459 C., which may be considered as exactly balancing the actual output.

It would be out of place to give a full description of the respiration calorimeter here. The general construction will be seen from the accompanying figure of the form of apparatus in use for patients in the Russell Sage Institute, New York. One of the most interesting details of its construction concerns the means taken to prevent any loss of heat from the calorimeter to the surrounding air. This is accomplished in the following way: The innermost layer of the wall is of copper; then, separated from this by an air space, is another wall of copper, outside of which are two wooden walls separated from each other and from the outer copper walls by air spaces. The two copper walls are connected through thermoelectric couples, so that an electric current is set up whenever there is any difference in their temperatures. The current is observed by means of a galvanometer placed outside the calorimeter, and from its movements the observer either heats up or cools down the outer copper walls so as to correct the difference of temperature causing the current. This is done by an electric heating device or by cold water tubes placed between the outermost copper and the innermost wooden walls. Since the temperature of the two copper walls is the same, there can be no exchange of heat between them, and consequently none of the heat that is absorbed by the inner copper walls is allowed to be carried away. All the heat given out by the animal is absorbed by the stream of cold water flowing through the coils

*A kilogram meter is the product of the load in kilograms multiplied by the distance in meters through which it is lifted.

of pipe in the chamber. The heat used to vaporize the moisture from skin and lungs must of course also be measured. This is done by collecting the water vapor in a sulphuric-acid bottle placed in the ventilating current. By multiplying the grams of water by the factor for the latent heat of vaporization, we obtain the calories of heat so eliminated.

"The calorimeter contains a comfortable bed and is provided with two windows, a shelf, a telephone, a fan, a light, and a Bowles stethoscope for counting the pulse. The ordinary experiment takes about as long as a trip from New York to New London. Patients, as a rule, doze from time to time or else try to work out some scheme by which they can amuse themselves without moving. After three or four hours they are rather bored by the quiet, and the observations are not prolonged beyond this time. They are allowed to turn over in bed once or twice an hour, but reading and telephoning are discouraged, since these increase the metabolism. The air in the box is fresh and pure, the patient suffers no discomfort, and objections to the procedure are very infrequent. Most of the patients are only too glad of the extra attention, and they insist that the calorimeter has a marked therapeutic value." (Du Bois.)

Normal Values.—Having thus satisfied ourselves as to the extreme accuracy of the method for measuring energy output, we shall now consider some of the conditions that control it. To study these we must first of all determine the *basal heat production*—that is, the smallest energy output that is compatible with health. This is ascertained by allowing a man to sleep in the calorimeter and then measuring his calorie output while he is still resting in bed in the morning, fifteen hours after the last meal. When the results thus obtained on a number of individuals are calculated so as to represent the calorie output per kilogram of body weight in each case, it will be found that 1 C. per kilo per hour is discharged —that is to say, the total energy expenditure in 24 hours in a man of 70 kilos, which is a good average weight, will be $70 \times 24 = 1{,}680$ C.

When food is taken the heat production rises, the increase over the basal heat production amounting for an ordinary diet to about 10 per cent. Besides being the ultimate source of all the body heat, food is therefore a direct stimulant of heat production. This *specific dynamic* action, as it is called, is not, however, the same for all groups of foodstuffs, being greatest for proteins and least for carbohydrates. Thus, if a starving animal kept at 33° C. is given protein with a calorie value which is equal to the calorie output during starvation, the calorie output will increase by 30 per cent, whereas with carbohydrates it will increase by only 6 per cent. Evidently, then, protein liberates much free heat during its assimilation in the animal body; it burns with a hotter flame than fats or carbohydrates, although before it is completely burned it may not yield

so much energy as is the case, for example, when fats are burned. This peculiar property of proteins accounts for their well-known heating qualities. It explains why protein composes so large a proportion of the diet of peoples living in cold regions, and why it is cut down in the diet of those who dwell near the tropics. Individuals maintained on a low protein diet may suffer intensely from cold.

If we add to the basal heat production of 1,680 C. another 168 C. (or 10 per cent) on account of food, the total 1,848 C. nevertheless falls far short of that which we know must be liberated when we calculate the available energy of the diet, which we may take as 2,500 C. What becomes of the extra fuel? The answer is that it is used for *muscular work.* Thus it has been found that if the observed person, instead of lying down in the calorimeter, is made to sit in a chair, the heat production is raised by 8 per cent, or if he performs such movements as would be necessary for ordinary work (writing at a desk) it may rise 29 per cent—that is to say, to 90 C. per hour. There is, however, practically no difference in the energy output of a person lying flat or lying in a semi-reclining position, as in a steamer chair. Allowing eight hours for sleep and sixteen hours for work, we can account for about 2,168 C., the remaining 300 odd C. that are required to bring the total to that which we know, from statistical tables of the diets of such workers, to be the actual daily expenditure, being due to the exercise of walking. If the exercise is more strenuous, still more calories will be expended; thus, to ascend a hill of 1,650 feet at the rate of 2.7 miles an hour requires 407 extra calories. Field workers may expend, in 24 hours, almost twice as many calories as those engaged in sedentary occupations.

Standard for Comparison

When the energy output per kilo *body weight* is determined in animals of varying size, the values are greater the lighter the animal. This is evident from the following results obtained on dogs:

Weight of dog	*Heat production in calories per kilo per day*
(1) 31.2	35.68
(2) 18.2	46.2
(3) 9.6	65.16
(4) 0.5	66.07
(5) 3.19	88.07

(Rubner)

When, on the other hand, instead of body weight, the area of the surface of the body is taken as the basis of calculation, results that are almost constant are obtained. Following are the results in the above animals on this basis:

Surface in square cm.	Heat production in calories per square meter of surface per day
(1) 10,750	1036
(2) 7,662	1097
(3) 5,286	1183
(4) 3,724	1153
(5) 2,423	1212

(Rubner)

Such results have prompted observers to conclude that the determining factor in the calorie output of warm-blooded animals is the *relative surface* of the animal. This is greater the smaller the animal, with the consequence that heat is more rapidly lost to the surrounding air from the surface, thus requiring more active combustion. Until quite recently it has been generally believed that such a relationship between body surface and heat production did actually exist, but, thanks to the work of F. G. Benedict[7] and E. F. and D. Du Bois[6], it is now known that the calculations were

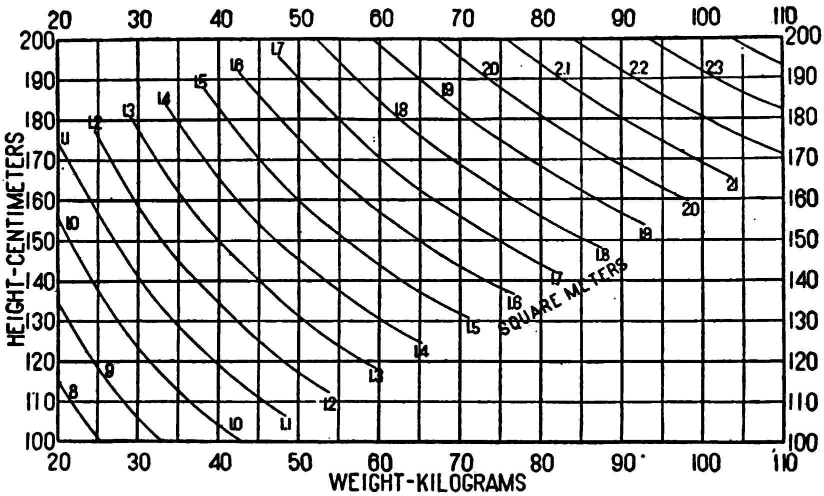


Fig. 175.—Chart for determining surface area of man in square meters from weight in kilograms (Wt.) and height in centimeters (Ht.) according to the formula: Area (Sq. Cm.) = Wt. 0.425 XHt. 0.725 X71.84. (From Dubois and Dubois, Arch. Int. Med., 1917, vol. 17.)

based upon incorrect computations of the body surface. In the older researches the calculation was made by using a formula known as Meeh's, in which weight was multiplied by a certain factor (viz., $12.312 \times \sqrt[3]{\text{weight}}$). Du Bois, however, has shown that an average error of 16 per cent is incurred in using this formula. For accurate measurement the body was covered with thin underwear, which was then impregnated with melted paraffin and reinforced with paper strips to prevent it from changing in area when removed. This model of the surface was afterwards cut up into flat pieces and photographed on paper of uniform thickness, the pat-

terns being then cut out, and weighed. From the results it was easy to calculate the actual surface area.

Where the height and weight are known, a fairly accurate computation of the surface can be secured by using the following formulas: $A = W^{0.425} \times H^{0.725} \times 71.84$; *A* being the surface area in square centimeters; *H* the height in centimeters; and *W*, the weight in kilograms. Based on this formula, a chart has been plotted from which the surface area may be determined at a glance (Fig. 175). Another method recently employed by Benedict is based on measurements made from photographs of the subject in various poses.

By the use of these more accurate measurements of body surface, it is now known that, although the surface-area law gives us constant results for the energy output of different individuals of similar build, and offers us a much more accurate basis for comparing those of different laboratory animals, yet it breaks down when applied to men in widely differing states of body nutrition. Thus, in the case of a man who starved for a month, the calorie output per square meter of surface decreased towards the end of the fast by 28 per cent. Obviously, therefore, it would be incorrect to draw conclusions regarding possible changes in energy output of a series of emaciated or corpulent individuals by comparison of their calorie output per square meter of surface with that of normal individuals.

The determining factor of energy output is undoubtedly the general condition of bodily nutrition—the active mass of protoplasm of the body (Benedict). That there is a relationship between the body surface and metabolism is undoubted, but the relationship is not a causal one. At present, therefore, the only safe method to employ in comparing the metabolism of normal and diseased individuals is that called by Benedict "the group method," in which the metabolism of groups of persons of like height and weight is compared, it being assumed that such individuals have the same general growth relations. For the application of this group method, however, more extensive data will be required than exist at present, and although some of the conclusions drawn from results computed on the surface-area basis may have to be revised, it is probable that they are in general correct.

Influence of Age and Sex

The energy output is low in the newly born; it increases rapidly during the first year, reaching a maximum at about three to six years of age, and then rapidly declining to about twenty, after which it declines much more slowly. The decline in the earlier years does not proceed steadily, however, for at the period just preceding the onset of puberty a decided increase becomes evident, indicating that at this period the metabolism of

the growing organism is being stimulated. Females have a lower energy output than males, and the stimulating influence of puberty is less marked in them.

In round numbers, 40 C. per square meter of surface per hour is the energy output of normal men, a 15 per cent deviation being considered as decidedly abnormal. The average metabolism of fat and thin subjects is the same, but that of women is 6.8 per cent lower than that of men. The basal metabolism of a group of men and women between the ages of forty and fifty was 4.3 per cent below the average for the larger group between the ages of twenty and fifty; and that of a group between fifty and sixty years was 11.3 per cent lower.

Influence of Diseases

The measurements have been made by the direct method which has just been described, but since the much simpler indirect method (page 554) yields comparable results, it is being adopted for clinical purposes. These results were obtained by making parallel determinations of energy output by both methods, in disease as well as in health. Some of the observations that have been made on the energy output in various diseases are as follows: In very severe cases of *exophthalmic goiter,* heat production may be increased by 75 per cent over the normal; in severe cases, by 50 per cent. The warmth of the skin and the sweating, which are prominent symptoms of this disease, are therefore accounted for by the increased elimination of heat, and it is considered possible that the other symptoms would be produced in any normal individual were his metabolism maintained for months or years at the high level which it occupies in goiter. In the opposite condition of *myxedema,* the energy output is markedly reduced, but rises slowly during treatment with thyroid extract, or much more rapidly with the very active thyroid hormone recently isolated by Kendall. In *diabetes* it has often been thought that the rapid emaciation and loss of strength were dependent upon an excited state of metabolism, or a useless burning up of the energy material. The most recent work, however, clearly shows that this is not the case, the basal metabolism as calculated per unit of body surface being within the limits indicated above. During the starvation treatment the energy output may be much below the normal. In uncompensated cases of *cardiorenal disease,* there is increased energy output. In *pernicious anemia* the metabolism is normal, although in severe cases there may be an increased demand for oxygen.

Even at the risk of repetition, it is important to point out that in all these diseases the energy output is the same whether measured directly or by the indirect method about to be described.

THE MATERIAL BALANCE OF THE BODY

We must distinguish between the balances of the organic and the inorganic foodstuffs. From a study of the former we shall gain information regarding the sources of the energy production whose behavior under various conditions we have just studied. From a study of the inorganic balance, although we shall learn nothing regarding energy exchange—for such substances can yield no energy—we shall become acquainted with several facts of extreme importance in the maintenance of nutrition and growth.

To draw up a balance sheet of organic intake and output requires an accurate chemical analysis of the food and of the excreta (urine and expired air).

Methods for Measuring Output

The principle by which the output is measured will be understood by referring to Fig. 176, from which it will be seen that the calorimeter is connected with a closed system of tubes provided with an air-tight ro-

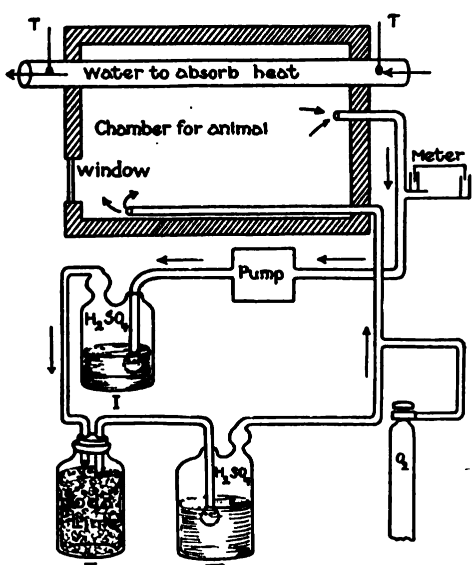


Fig. 176.—Diagram of Atwater-Benedict respiration calorimeter. As the animal uses up the O_2, the total volume of air shrinks. This shrinkage is indicated by the meter, and a corresponding amount of O_2 is delivered from the weighed O_2-cylinder. The increase in weight of bottles II and III gives the CO_2; that of I, the water vapor.

tary blower or pump to maintain a constant current of air, as indicated by the arrows. Following the air stream as it leaves the chamber, we note a side tube connecting with a meter to indicate changes in volume of the

air in the system. Beyond this and the pump is a specially constructed bottle containing concentrated H_2SO_4, then one containing soda lime, and lastly another H_2SO_4 bottle. The first H_2SO_4 bottle absorbs all the water vapor contained in the air coming from the chamber; the soda lime bottle absorbs the CO_2, and the second H_2SO_4 bottle absorbs water that is produced in the chemical reaction involved in the absorption of the CO_2 by the soda lime ($2NaOH+CO_2=H_2O+Na_2CO_3$). By weighing these absorption bottles before and after an animal has been for some time in the chamber, the weight of H_2O and of CO_2 given out can be determined. Another side tube leads to an oxygen cylinder, the valve of which is manipulated so as to cause oxygen to be discharged into the system at such a rate as to compensate exactly for that used up by the animal, as indicated by the behavior of the meter. The amount of oxygen required is determined either by weighing the oxygen cylinder before and after the observation or by measuring the volume of oxygen used by passing it through a carefully calibrated and very sensitive water meter inserted on the side tube that connects the O_2 cylinder with the main tubing of the system. Since muscular activity causes pronounced changes in the rate of metabolism, means are usually taken to secure graphic records of any movements made during the observation.

The growing importance in clinical investigations of measurements of the respiratory exchange and the necessity for having methods that are as simple as is consistent with accuracy, have led to the introduction of several other forms of apparatus, of which those of F. G. Benedict and of Tissot* are the most important. In the former a tightly fitting mask, applied over the nose and mouth is connected, by a short T-piece, with the same tubing as that used in the respiration calorimeter. The patient thus breathes in and out of the air stream that is passing along the tubing without any of the obstruction experienced when the breathing has to be performed through valves, as in the older (Zuntz) forms of portable respiratory apparatus. It is particularly for studies on man that this apparatus has been devised. The Tissot and Douglas methods are shown in Figs. 179 and 180.*

To complete the investigation, it is necessary that the urine and feces be collected and the nitrogen excretion measured. When the respiratory excreta are measured over a considerable period of time, as in the large calorimeter, the urine is collected for the same period, but when shorter respiratory measurements are made, the urine of the twenty-four hours is usually taken.

Principles Involved in Calculating the Results.—Provided with the analyses furnished by the above methods, we proceed to ascertain the total

*The Tissot method will be found described in full elsewhere (page 554).

amounts of nitrogen and carbon excreted and to calculate from the known composition of protein how much protein must have undergone metabolism. We then compute how much carbon this quantity of protein would account for, and we deduct this from the total carbon excretion. The remainder of carbon must have come from the metabolism of fats and carbohydrates, and although we can not tell exactly which, yet we can arrive at a close approximation by observing the respiratory quotient (R. Q.), which is the ratio of the volume of carbon dioxide exhaled to that of oxygen retained by the body in a given time, i. e., $\frac{CO_2}{O_2}$. By observing this quotient, therefore, we can approximately determine the source from which the nonprotein carbon-excretion is derived.

Having in the above manner computed how much of each of the proximate principles has undergone metabolism, we next proceed to compare intake and output with a view to finding whether there is an equilibrium between the two, or whether retention or loss is occurring.

It may serve to make clear the methods by which these calculations are made to study the following example:

Example of a Metabolism Investigation.—It is desired to know whether a diet containing 125 grams protein, 50 grams fat, and 500 grams carbohydrate is sufficient for a man doing a moderate amount of work.

INTAKE

	Carbon	Nitrogen	Calories
Protein,	62 gm.	20 gm.	512.5
Carbohydrate,	200	—	2050.0
Fat,	38	—	465.0
Total,	300 gm.	20 gm.	3027.5

OUTPUT

	Carbon		Nitrogen
In urine,	11 gm.	(16.5×0.67)	16.5 gm.
In feces,	5		1.0
In the breath,	254		—
Total,	270 gm.		17.5 gm.

Retained in Body.—30 gm. carbon and 2.5 gm. nitrogen. This amount of nitrogen represents $2.5 \times 6.25 = 15.6$ gm. *protein* or 75 gm. *muscle.* Now, this amount of protein will account for 8.25 gm. carbon; so that $30 - 8.25 = 21.75$ gm. carbon represents $21.75 \times 1.3 = 28.3$ gm. *fat.* On this diet, therefore, the subject retains in his tissues 15.6 gm. protein and 28.3 gm. fat per diem.

Furnished with these data we may now proceed to compute how much energy must have been liberated in the body.

To express the above result in terms of energy liberated, we know that 3027.5 C. were supplied and that all these have been used except $15.6 \times 4.1 = 64$ retained as protein, and $28.3 \times 9.3 = 263.2$ retained as fat; or *in toto* 327.2 C. We find, therefore, that $3027.5 - 327.2 = 2{,}700$ C. have been required.

This is called the method of *indirect calorimetry,* and it has been clearly established by numerous observations that the results agree exactly with those secured by the method of *direct calorimetry* described above. For most purposes the indirect method is quite satisfactory, and it is especially valuable in cases in which there are considerable and sudden changes in body temperature. That the results by the two methods should agree shows clearly that the law of the conservation of energy must apply in the animal body, for it is evident that if any energy were derived from outside the body other than that taken with the food, the results by the direct method would be higher than those by the indirect.

CHAPTER LXI

THE CARBON BALANCE

Before proceeding to discuss the special metabolism of proteins, fats and carbohydrates, it will be advantageous to consider briefly some general facts concerning the excretion of carbon dioxide and the intake of oxygen. In the first place, it is important to note that the *extent* of the combustion process in the animal body is proportional to the amount of oxygen absorbed and of carbon dioxide produced, whereas the *nature* of the combustion is indicated by the ratio existing between the amounts of carbon dioxide expired and of oxygen retained in the body. An investigation of the carbon balance, in other words, is partly quantitative and partly qualitative—quantitative in the sense that it indicates how intensely the body furnaces are burning, and qualitative in the sense that it tells us what sort of material is being burned at the time.

THE RESPIRATORY QUOTIENT

Influence of Diet.—The respiratory quotient is determined by comparison of the volume of carbon dioxide expired with the volume of oxygen meanwhile retained in the body or, as a formula,

$$\frac{\text{Vol. } CO_2 \text{ expired}}{\text{Vol. } O_2 \text{ retained}}.$$

For the sake of brevity the respiratory quotient is often written R. Q. That it serves as an indicator of the kind of combustion occurring will be evident from the following equations:

1. *Carbohydrate:* $C_6H_{12}O_6 + 6O_2 = 6CO_2 + 6H_2O$
(Dextrose.)

$$\therefore \text{R.Q.} = \frac{CO_2}{O_2} = \frac{6}{6} = 1.$$

2. ***Fat:*** $C_3H_5(C_{18}H_{33}O_2)_3 + 80O_2 = 57CO_2 + 52H_2O$
(Olein.)

$$\therefore \text{R.Q.} = \frac{CO_2}{O_2} = \frac{57}{80} = 0.71$$

3. ***Protein:*** $C_{72}H_{112}N_{18}O_{22}S + 77O_2 = 63CO_2 + 38H_2O + 9CO(NH_2)_2 + SO_3$
[Empirical formula for albumin (Lieberkühn).]

$$\therefore \text{R.Q.} = \frac{CO_2}{O_2} = \frac{63}{77} = 0.82$$

4. *Conversion of fat into carbohydrate:*

$$2C_3H_5(C_{18}H_{33}O_2)_3 + 64O_2 = 16C_6H_{12}O_6 + 18CO_2 + 8H_2O$$

(Olein.)

$$\therefore \text{ R.Q.} = \frac{CO_2}{O_2} = \frac{18}{64} = 0.281$$

5. *Conversion of carbohydrate into a mixed fat:*

$$13C_6H_{12}O_6 = C_{55}H_{104}O_6 + 23CO_2 + 26H_2O.$$

(Oleostearopalmitin.)

Taking carbohydrates first, the general formula may be written CH_2O, from which it is plain that, to oxidize the molecule, oxygen will be required to combine with the carbon alone, according to the equation, $CH_2O + O_2 = CO_2 + H_2O$. In other words, the volume of carbon dioxide produced by the combustion will be exactly equal to the volume of oxygen used in this process, in obedience to the well-known gas law that equimolecular quantities of different gases occupy the same volume. The respiratory quotient is therefore unity (Equation 1). With fats and proteins, however, the general formula must be written $CH_2 + O$, indicating therefore that for its complete oxidation the molecule must be supplied with oxygen in sufficient amount to combine not only with all of the carbon, but also with some of the hydrogen, forming water; so that the volume of CO_2 produced will be *less* than the volume of oxygen retained, and the respiratory quotient will be less than unity. As a matter of fact, as the above equations show (2 and 3), the respiratory quotient for fats and proteins lies somewhere between 0.7 and 0.8, being usually nearer 0.7 in the case of fats, and nearer to 0.8 in the case of proteins.

That the conditions hypothecated in the equations exist in the animal body during the combustion of the foodstuffs can easily be shown by observing the respiratory quotient of animals on different diets. An herbivorous animal, such as a rabbit, when it is well fed gives invariably a respiratory quotient of about 1, whereas a strictly carnivorous animal, such as the cat, gives a respiratory quotient of about 0.7. Even more striking perhaps is the comparison of the respiratory quotients in an herbivorous animal while it is well fed and after it has been starved for a day or two. In the latter case the respiratory quotient will fall to a low level because, by starvation, the animal has been compelled to change its combustion material from the carbohydrate of its food to the protein and fat of its own tissues.

As already explained (page 545), it is from the respiratory quotient that we are enabled to tell what proportions of fat and carbohydrate, respectively, are undergoing metabolism. A useful table showing the percentage of calories produced by each of these foodstuffs, after allowing for protein, is given by Graham Lusk (see page 565).

Influence of Metabolism.—Apart from diet, the respiratory quotient may often be altered by changes in the metabolic habits of the animal. These are most conspicuously exhibited in the case of hibernating animals. In the autumn months, when the animal is eating voraciously of all kinds of carbohydrate food and depositing large quantities of adipose tissue in his body, the respiratory quotient may be considerably greater than unity, indicating therefore either that relatively more carbon dioxide is being discharged or less oxygen retained. As a matter of fact, it can easily be shown that it is the former of the causes that is responsible for the higher quotient, the explanation for the increased production of CO_2 being that, as the carbohydrate changes into fat, the relative excess of carbon in the former is got rid of as CO_2, as indicated in Equation 5. On the other hand, if the animal is examined while in his winter sleep, it will be found that the respiratory quotient is now extremely low, often not more than 0.3 to 0.4, which may be interpreted as indicating either an excessive absorption of oxygen or a markedly decreased excretion of carbon dioxide. As a matter of fact, there is a great diminution in both the excretion of carbon dioxide and the intake of O_2, because the whole metabolic activity of the animal is extremely depressed, but this diminution affects the oxygen to a much less degree, indicating therefore a relative increase in the oxygen retention. The explanation is that the oxygen is being used in the chemical process involved in the conversion of the fat back into carbohydrate.

Whatever may be the relationship between fat and carbohydrate in the nonhibernating animal, there is no doubt that during hibernation, before the fat stores are burned, fat is converted into something closely related to carbohydrates, the equation for the process being represented as given above (No. 4).

In man and the higher mammalia, the only condition apart from diet which can affect the nature of the combustion process is disease; thus in total diabetes (page 678) the organism loses the power of burning carbohydrate, so that whatever the diet may be, the respiratory quotient is very low, never higher than that representing combustion of fat and protein. It has been claimed by certain investigators that in diabetes the respiratory quotient may fall considerably below 0.7, indicating, as in hibernating animals, that fat is being converted into carbohydrate. The most recent and carefully controlled observations, however, deny this claim, and for the present we must assume that in the body of man fat is not converted into carbohydrate (see page 664). In numerous other diseases investigated by Du Bois and others[6] no qualitative change in the combustion processes in man has been brought to light.

THE MAGNITUDE OF THE RESPIRATORY EXCHANGE

It is evident that the amount of carbon dioxide expired and of oxygen retained will be proportional to the energy liberation in the animal body. Even at the risk of repetition it should be noted that the energy exchange can be very accurately calculated from the result of the material balance sheet—indirect calorimetry, as it is called (page 562). On account of the comparative simplicity of measuring the carbon dioxide output and oxygen intake, it is natural that many of the observations that have been made on energy production in the animal body depend on the use of this method, justification for which is found in the complete agreement between the results of direct and indirect calorimetry in a great variety of diseases and conditions in man (Du Bois[6]).*

In the first place, it is interesting to compare the respiratory exchanges of different animals computed per kilo body weight. This is shown in the following table.

ANIMAL	WEIGHT GM.	OXYGEN ABSORBED PER KILO AND HOUR GM.	CARBON DIOXIDE DISCHARGED PER KILO AND HOUR GM.	$\frac{VOL.\ CO_2}{VOL.\ O_2}$	TEMPERATURE OF AIR
Insecta					
Field cricket	0.25	——	2.305	—	—
Amphibia					
Edible frog		0.063 (44.2 c.c.)	0.060 (30.76 c.c.)	0.69	15°-19°
		0.105 (73.4 c.c.)	0.1134 (57.7 c.c.)	0.78	—
Aves					
Common hen	1280	1.058 (740 c.c.)	1.327 (675 c.c.)	0.91	19°
Pigeon	232-380	——	3.236	—	—
Sparrow	22	9.595 (6710 c.c.)	10.492 (5334.5 c.c.)	0.79	18°
Mammalia					
Ox	638,000 660,000	——	0.389-0.485	—	—
Sheep	66,000	0.490 (343 c.c.)	0.671 (341 c.c.)	0.99	16°
Dog	6213	1.303 (911 c.c.)	1.325 (674 c.c.)	0.74	15°
Cat	2464	1.356 (947 c.c.)	1.397 (710 c.c.)	0.75	-3.2°
	3047 ,,	0.645 (450 c.c.)	0.766 (389 c.c.)	0.86	29.6°
Rabbit	1433	1.012	1.354	0.97	18°-20°
Guinea pig	444.9	1.478	1.758	0.86	22°
Rat (white)	80.5	——	3.518 (1789 c.c.)	—	7°
Mouse "	25	——	8.4	—	17°
Man	66,70	0.292	0.327	—	—

(Modified from Pembrey.)[17]

*For the convenience of those who may desire to know more about the methods of analysis that are suitable in the clinic, a chapter on the subject will be found beginning on page 554.

Several factors operate to explain these differences, and of these the following are of importance:

1. **The Body Temperature.**—Increase in body temperature entails increased combustion. This explains why the metabolism of a bird is greater than that of a mammal of the same size, for, as is well known, the temperature of a bird is two or three degrees centigrade above that of other animals. Rise in body temperature also explains, in part at least, the increased metabolism observed in fever.

2. **The Temperature of the Environment.**—In considering this we must distinguish between the effect produced on warm-blooded and on cold-blooded animals. Since the body temperature of a cold-blooded animal is only one or two degrees Centigrade above that of its environment, it follows that the metabolic activity will be directly proportional to the temperature of the latter. In a warm-blooded animal, on the other hand, the body temperature remains constant whatever changes may occur in that of the environment, this constancy of body temperature being dependent on the fact that the intensity of the combustion processes is inversely proportional to the cooling effect of the atmosphere. Thus, suppose the external temperature should fall, then the loss of heat from the body will tend to become greater, and to maintain the body temperature at a constant level, the body furnaces must burn more briskly, with the result that an increased excretion of carbon dioxide and intake of oxygen will occur.

This influence of the surrounding atmosphere on the metabolic activity of warm-blooded animals has, as already pointed out, been used by several investigators to explain the greater combustion per kilo body weight of small as compared with large animals. The argument is that, since the surface of small animals relatively to their mass is much greater than in large animals, the cooling of the small animals will be proportionately greater. The relationship between surface and mass is shown by taking two cubes and putting them together; the mass of the two cubes is equal to double that of either cube, whereas the surface is less than double, since two aspects of the cubes have been brought together. To prove the contention, the respiratory exchange has been computed per square meter of surface instead of per kilo body weight, with the result that a very close correspondence in the metabolism of different animals has been observed; but this question has already been discussed, and we now know that the law of cooling can not be the only one that determines extent of the respiratory exchange (see page 541).

3. **Muscular Exercise.**—This has a most important influence on the exchange and it is particularly in connection with it that studies in carbon-dioxide output and oxygen intake have been of great practical value, par-

ticularly when the investigations are undertaken on men doing ordinary types of muscular exercise, such as walking or climbing. It is true that the influence of muscular exercise on the energy metabolism may also be studied by having a person in the calorimeter do exercises on an ergometer, but the results thus obtained are in many ways not nearly so valuable as those which can be secured by observing the respiratory exchange of persons doing ordinary types of muscular exercise in the open. The following table of observations on horses is of interest in this connection.

CONDITION	AIR EXPIRED IN LITERS PER MINUTE	CARBON DIOXIDE DISCHARGED IN LITERS PER MINUTE	OXYGEN ABSORBED IN LITERS PER MINUTE	$\frac{CO_2}{O_2}$
Rest	44	1.478	1.601	0.92
Walk	177	4.342	4.766	0.90
Trot	333	7.516	8.093	0.93

It will be observed that the metabolism increases extraordinarily for even a moderate degree of work, but that at the same time the respiratory quotient remains constant. From observations on the respiratory exchange of working men and animals, extremely important facts concerning the efficiency of muscular work have been secured. The form of respiratory apparatus (Zuntz or Douglas) employed for this purpose must be capable of being strapped on the man's back without causing any embarrassment to his bodily movements. By a comparison of the respiratory exchange with the amount of work done, the efficiency of the work can readily be determined. It has been found, for example, that the efficiency is much greater after the man or animal has got into the swing of the work, his energy expenditure per unit of work being much greater during the first half hour's work in the morning than it is later on. This indicates that after a little practice the muscles can execute a given movement and perform a given amount of work much more smoothly than when they are not in training. Another interesting outcome of the investigations has been to show that work done under abnormal conditions that tend to produce any kind of muscular strain is done inefficiently. It has been found in marching soldiers, for example, that the slightest abrasion of the foot greatly increases the energy expenditure, for the man, in trying to avoid the pain produced by the abrasion, brings into operation muscular groups that are really not required for the efficient performance of the movement, but are used instead to avoid pressure on the sore. Fatigue also causes inefficient performance of work; that is to say, the fatigued person, on attempting

the same amount of work as he performed before becoming fatigued, will do so at a much greater expenditure of energy.

There is a diurnal variation in the respiratory exchange, which is in general parallel with the body temperature; it rises during the day, the time of activity and work, and falls during the night, the time of rest and sleep. Food also affects respiratory exchange, but it will be unnecessary to go into this further after what has been said on page 547.

CHAPTER LXII*

A CLINICAL METHOD FOR DETERMINING THE RESPIRATORY EXCHANGE IN MAN

By R. G. Pearce, B.A., M.D.

Principle.—Since the determination of the respiratory exchange in man is of some importance in the study of certain diseases of the respiration, circulation and metabolism, and also because directions for carrying out the necessary procedures are not generally available, we have thought it might be of assistance to include here brief directions for the Tissot and the Douglas methods. These methods have been found to compare favorably in accuracy with others in use at present,† and because of their adaptability and simplicity they are specially suited for clinical work.

By these methods the energy metabolism of the body is calculated from oxygen consumption or carbon dioxide excretion per minute (indirect calorimetry) (page 546), the figures for which are determined from the volume and percentile gaseous composition of the expired air.

The subject breathes through valves which automatically partition the inspired and expired air. The expirations from a number of respirations are collected in a spirometer or bag, and the volume of the respirations per minute is determined. The gaseous composition of the expired air is determined by gas analysis, and the oxygen consumption and energy output of the body are calculated from the data obtained.

Description and Use of Parts of the Apparatus: 1. The Mouthpiece and Valves.—The mouthpiece is made of soft pure gum rubber, and consists of an elliptical rubber flange having a hole in the center 2 cm. in diameter, to which on one side a short rubber tube is attached. On the opposite side of the hole, at right angles to the rubber flange, are attached two rubber lugs. The rubber flange is placed between the lips, and the lugs are held by the teeth. The rubber tube of the mouthpiece is connected to the tube carrying the valves. The nose must be tightly closed if mouth breathing is used. This is accomplished by a nose clip, which consists of a V-shaped metal spring, the ends of which are provided with felt pads. A toothed rachet is attached to the ends of the

*This chapter is added for the convenience of workers in this subject.
†Carpenter: Carnegie Institution of Washington Reports, No. 216, 1915.

spring, and serves to hold the spring tightly clamped on the nostrils in the proper position (see Fig. 177).

Some individuals experience great distress when made to breathe through the mouth. For these it is best to use a face mask. Unfortunately at the present time no mask is entirely satisfactory. Perhaps the best is one sold by Siebe, Gorman & Co.,* which is pictured in the cut.

Fig. 177.—*A*, Nose clip; *B*, Face mask; *C*, Mouth piece.

After being placed in position the face mask should be tested for leaks, which can be done by putting soap around the edges.

2. The Valves.—The valves of Tissot are probably the best for the purpose, but they are expensive and difficult to obtain. We have made perfectly satisfactory valves from the prepared casings used in the manufacture of bologna sausage. These can be obtained preserved in salt, and they will keep indefinitely on ice. When needed a short piece

*This mask has been used extensively by Carpenter. The agent in this country is H. N. Elmer, 1140 Monadnock Bldg., Chicago.

is taken, washed free from salt by allowing water from the tap to run through it, and softened in a weak glycerine solution. The gut becomes very soft and pliable, and does not dry quickly. A piece of the casing about 10 cm. long is threaded through a glass tube of about 15 mm. bore and 4 to 6 cm. long. One end of the casing is brought around the outside of the tubing and secured by means of a thread. The lower end of the membrane is pinched off and the casing is then cut a little more than half way across its middle, so that the opening will lie just within the free end of the tube when the casing is drawn back through it. The loose end of the casing is slightly twisted—an essential procedure—and is then secured by a thread on the outer side of the tube. If properly made, the valve will work freely without vibration, and the opening be sufficiently large to allow a good current of air to pass. It should collapse instantly and be air-tight when the current of air is reversed. The back lash, or lag of closure, of these valves is extremely small, and they will open or close with a pressure of air not exceeding the pressure

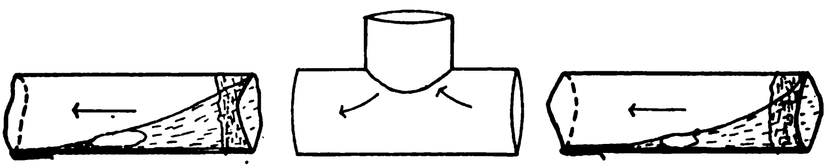


Fig. 178.—Diagram of respiratory valves.

changes in normal respiration. When not in use, the valves should be kept in glycerine water on ice. Valves prepared in this way have been in use a month without loss of efficiency. They are, however, made with so great ease that new valves are provided for each subject, and they are therefore especially adapted to ward work (Fig. 178).

The valves are inserted in reverse order into a supporting metal T-piece, and the joints made air-tight by tape. The stem of the T is connected with the mouthpiece. Through a rubber tube of about 3/4 inch bore, the expired air is collected in the spirometer, or Douglas Bag.

3. The Tissot Spirometer is pictured in Fig. 179. We have found the 100-liter size to be very serviceable in the clinic. This instrument is mounted on a platform having rubber wheels, and can be moved about the wards with ease. The bell of the spirometer is made of aluminum and is suspended in a water-bath between the double walls of a hollow cylinder made of galvanized iron. The height of the bell is 72 cm. and the diameter 42 cm. An opening at the bottom of the cylinder connects through a three-way stopcock with the rubber tube leading from the expiratory valve of the mouthpiece (see Fig. 177).

The bell is counterpoised by means of a weight. In the original Tissot spirometer an automatic adjustment permitted water in amount equal to the water displaced by the bell to flow from the spirometer cylinder into a counterpoise cylinder as the bell ascended out of the water.

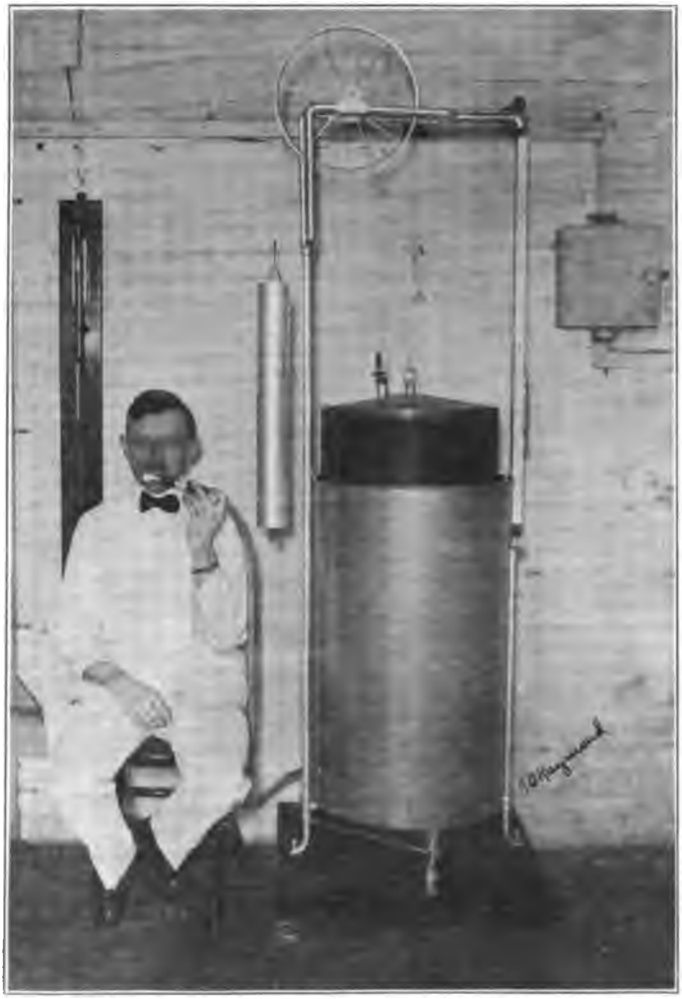


Fig. 179.—The Tissot spirometer. In actual experiment, subject is reclining or lying down and the valves and mouthpiece are held with a clamp.

The bell, being heavier out of water than when it is immersed, is accordingly counterpoised in any position, although Carpenter has shown that this refinement is unnecessary. An opening in the top of the spirometer permits the insertion of a rubber stopper, through which are passed a thermometer, a water manometer, and a stopcock with tube for drawing

the sample of air. A scale on the side of the instrument gives the volume of the air.

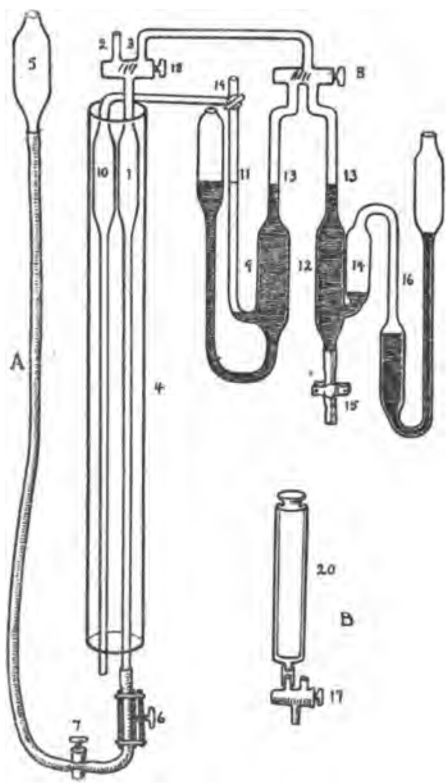
During an observation the subject sits in a reclining position or lies upon a couch. When the bell of the spirometer is placed at zero, the mouthpiece adjusted in the mouth, and the nose clamped, respiration is started, the expirations being passed through the stopcock, which is so turned as to allow them to pass to the outside air. After a few minutes the stopcock is turned so that the expirations are passed into

Fig. 180.—The Douglas bag method for determining the respiratory exchange. The arrangement of mouthpiece, valves, and connecting tubes shown here has been found to be more convenient than that recommended by Douglas.

the spirometer for a definite length of time. At the end of the period the cock is again turned, and after the barometric pressure, temperature, and volume of the air have been noted, the composition of the air is determined in the Haldane gas analysis apparatus.

4. The Douglas Bag.—The Douglas bag is made of rubber-lined cloth, and is capable of holding from 50 to 100 liters. It is especially useful for investigations during exercise, since it is fitted with straps so that the bag can be fastened to the shoulders (Fig. 180). It is then connected with the valves, the mouthpiece of which is placed between the lips.

Respirations are commenced with the three-way valve turned so as to allow the expirations to pass directly outside. After respiratory equilibrium is established, the three-way valve is turned during an inspiratory period so that the succeeding expirations may pass into the bag. The time required to fill the bag comfortably is determined with a stopwatch. The air which has been collected in the bag during the period is thoroughly mixed and passed through a meter, the temperature and barometric pressure are noted, and a sample analyzed in the Haldane



A. *B.*

Fig. 181.—Haldane gas apparatus (*A*) and Pearce sampling tube (*B*).

gas-apparatus. The bag should be emptied completely by rolling it up when nearly empty.

5. **The Haldane Gas-analysis Apparatus.** PRINCIPLE.—The Haldane method of analysis of expired air is simple and easily learned. The apparatus (Fig. 181) consists of a gas burette, a control burette of the same size (both surrounded with a water jacket), and bulbs containing dilute caustic potash or soda solution for the absorption of the carbon dioxide and an alkaline pyrogallate solution for the absorption of the

oxygen. The gas burette is connected with the bulbs by a two-way stopcock, which allows a sample of gas to pass into either bulb. A control tube (*10*) is put into connection with the burette through a manometer tube, which is connected with the alkali bulb, and can be made to compensate for any changes in temperature that may occur during the course of the analysis. For an analysis the gas is transferred to the burette from the sampling tube, saturated with water vapor over mercury, and then measured, after which it is transferred into the caustic solution to free it from CO_2, and returned to the burette to determine the loss of volume due to CO_2 absorption. It is then transferred into the alkaline pyrogallate solution, which frees it from oxygen, after which it is again brought back to the burette to determine the loss in volume due to the absorption of the oxygen.

THE APPARATUS.—The detail of the Haldane apparatus is shown in the accompanying cut. The measuring burette (*1*) holds 21 c.c. The bulb is of 15 c.c. capacity, and the graduated stem, which is about 4 mm. in bore and 60 cm. in length, is graduated to 0.01 c.c. from 15 c.c. to 21 c.c. The stopcock at the top of the burette is double-bored, so that in one position air can be drawn in from a gas sampler (*2*) and in another sent into the absorption bulbs (*3*). The lower part of the burette extends through the rubber cork at the bottom of the water jacket (*4*). A piece of rubber tubing is attached to the bottom of the burette and is passed through a metal tube, furnished on its inside with a metal disc which presses against the rubber tubing, the pressure being controlled by means of a fine adjusting screw (*6*). Below this a glass stopcock (*7*) connects with rubber tubing to the mercury leveling bulb (*5*). The absorption bulb for CO_2, containing 20 per cent NaOH or KOH (*9*), is put in connection with the burette by suitably turning stopcocks (*3* and *8*).* The control burette (*10*) is also in connection with this bulb through the manometer tube (*11*).† Any variation in temperature which may occur during the analysis will cause the level of the alkaline solution in the manometer to change.

When final readings of the shrinkage of volume are made, the level of the caustic solution is returned to the level of that in the manometer. By so doing any error due to temperature changes is avoided, since change in temperature must be equal in the two burettes.

The absorption bulb for oxygen (*12*) is filled with a solution made by dissolving 10 grams of pyogallic acid in 100 c.c. of a nearly saturated KOH solution. The specific gravity of the KOH should be 1.55, which is obtained approximately by dissolving the sticks (pure by alcohol) in an

*The stopcock (*8*) is double-bored, so that the tube leading from the burette can be brought into connection with either 9 or 12.

†This tube also has a three-way stopcock (*19*), so that it may be opened to the outside.

equal weight of water. The mark (*13*) on the stem of the bulb indicates the level at which the solutions should stand. Enough pyrogallate solution is introduced through tube *15* to fill bulbs *12* and *14* two-thirds full. Then pyrogallate solution is poured into tube *16* until the difference in level of the fluids is sufficient to produce enough pressure to raise the level of the pyrogallate solution in *12* to the level *13* on the stem. Stopcock *8* must be open during this procedure. It may be necessary to add or take away a little pyrogallate solution through *15* to attain the above level.

Care must be taken to allow for complete absorption of oxygen from the air that is entrapped between *14* and *16* before an analysis is made; otherwise changes will be produced in the level of the pyrogallate solution. The air in the capillary tubing connecting the burettes with the absorption bulbs must also be freed of CO_2 and O_2. This can be accomplished by making a dummy analysis of atmospheric air before the real analysis. Great care must be taken to have atmospheric pressure in all the tubes at the start of the analysis. This is accomplished by opening the stopcock in the burette first to atmospheric air and then to the absorption bulbs, until no further change in the level of the fluids in the stems of the absorption bulbs occurs. This level is then marked and used as the standard. A small amount of water in the burette over the mercury assures saturation of the air with water vapor. Time for drainage must be allowed before making readings.

A very serviceable *sampling tube* for the transfer of air can be made from a 30 c.c. ground-glass syringe, to which is attached a two-way stopcock. A cut of this is shown in Fig. 181. The dead space in these syringes is washed out by working the piston back and forth several times. A thin coating of vaseline prevents leakage of the gas. We have found that these sampling tubes will retain a sample of expired air without change up to eight hours.

Manipulation of Apparatus.—The sampling syringe (*20*) is attached to opening *2* of the burette, and its stopcock (*17*) opened to atmospheric air. The level of the mercury is raised to the level of the stopcock of the syringe and is then turned so that syringe and burette are in communication. The bulb of mercury is lowered so that the mercury falls in the burette. This draws the piston of the syringe with it, and fills the burette with air from the syringe. It is advisable to put a little positive pressure on the piston of the syringe in the maneuver to prevent possible leakage. When all of the air is in the burette a slight positive pressure is produced in the burette by gently pressing on the piston, and immediately thereafter the stopcock on the syringe (*17*) is again turned to the original position. This allows the pressure of air in the burette to come to that

of the atmosphere. The height of the mercury is now adjusted to a convenient height in the burette by closing cock *7* and turning the milled screw *6*. The cock *18* is now made to communicate with the absorption bulbs. If the air in the burette is at atmospheric pressure, no change will occur in the level of the fluids. The reading is then taken on the burette.

The next step in the analysis consists in turning stopcock *8* to communicate with the caustic soda solution in bulb *9*, and the leveling tube (*5*) is raised, forcing mercury into the burette and the air into bulb *9*. The gas is passed back and forth several times until absorption is complete, as can be determined by the fact that the level of the mercury in the burette remains constant when the fluid in the bulb is returned to its original level (*13*) on the stem. In this adjustment it is convenient to make the gross leveling by the mercury bulb and the fine leveling by closing *7* and turning *6* until the fluid in *9* is at the original height. The reading on the burette indicates the loss in volume due to the CO_2 absorbed.

The oxygen is removed by a similar procedure, the gas being passed into the alkaline pyrogallate solution by turning cock *8* to communicate with bulb *12*. The absorption of oxygen is slower than for CO_2, and more care must be taken to get complete absorption. The air in the tubing between the fluid in *9* and stopcock *8* must be washed out several times in order to get the oxygen which is left in it after the absorption of the CO_2. When this is complete, the final reading on the burette is made and the loss in volume from the second reading represents the oxygen.

THE CALCULATIONS

The calculation of the percentile composition of the air and of the respiratory quotient is represented in the following example of an actual analysis:

(The temperature and barometric pressure as taken at the time of the experiment were 20° C. and 747 mm. Hg.)

CO_2 analysis—

1st reading of burette	20.00
2nd reading of burette after absorption of CO_2	19.20
CO_2 absorbed	0.80

$0.80 \div 20 = 4.0$ per cent CO_2 in expired air.

O_2 analysis—

2nd reading of burette	19.20
3rd reading of burette after absorption of O_2	15.90
O_2 absorbed	3.30

$3.30 \div 20 = 16.50$ per cent of O_2 in expired air.

Determination of R.Q.—

O_2 in atmospheric air = 20.94%

$O_2 - CO_2$ in expired air $(16.50 + 4) = 20.50\%$

$100 - 20.94 = 79.06\%$,* N in atmospheric air.

$100 - 20.50 = 79.50\%$, N in expired air.

Since the nitrogen is not changed in volume, the last figure shows that more oxygen must have been taken in during inspiration than $O_2 + CO_2$ has been given back in expiration. This obviously must be taken into account in the calculations. The amount of O_2 actually inspired for each 100 c.c. of air expired is found as follows:

$\frac{20.94 \text{ (\% } O_2 \text{ in atmospheric air)}}{79.06 \text{ (\% } N_2 \text{ in atmospheric air)}} \times 79.50$ (% N_2 in expired air); or 0.265 (constant factor $\times$ 79.5 (% N found for this observation) = 21.07, the volume of O_2 which would have been present in expired air to account for N present.†

$21.07 - 16.50 = 4.57\%$ O_2 actually absorbed.

$4.00 - 0.03$ (CO_2 in inspired air) $= 3.97\%$ CO_2 excreted.

$\therefore \frac{3.97}{4.57} = 0.87$, the respiratory quotient, or ratio of CO_2 excreted to O_2 absorbed.

Total Gas Exchange.—The volume of air expired in 15 minutes into the Tissot spirometer was found to be 100 liters measured at 20° C. and 747 mm. Hg (brass-scale barometer). This volume of gas must be corrected so as to give the volume of dry air at 0° and 760 mm. Hg. To do this two things must be taken into account. (1) Since the expired air is saturated with water, the pressure due to water vapor must be subtracted from the observed barometric pressure to obtain the true pressure. The vapor tension of water for various temperatures is given in Table II on page 564. (2) The barometer tube lengthens or contracts with heat or cold, and therefore the barometric readings must be corrected. The corrections for ordinary barometric readings are found in Table III, page 565. The figure corresponding to the temperatures is subtracted from the barometric reading in order to obtain correct barometric pressure.

In the above experiment, the correction for the barometer is 2.41 mm. (see Table III, page 565), and that for vapor tension at 20° C. is 17.4 (see Table II, page 564).

Actual Barometric Pressure.—$747 - (17.5 + 2.39) = 727.21$ mm. The coefficient of expansion of gases is taken as 0.003665) or 1/273; therefore the volume of 0° equals the volume at 1° divided by $1 - 0.003665$ t; and hence

*This is the constant O percentage in air.

†This calculation can be simplified by using an abbreviated table (page 564) giving the O_2 figure corresponding to the various percentages of N in the expired air.

$$Vo = \frac{V \times 273}{273 + t} = \frac{V}{1 + 0.003665\ t}$$, when Vo = Volume at 0° and V = Volume at t°.

The volume of gas being inversely as the pressure, $Vo = \frac{VP}{760}$, where V = volume at P pressure; or working both corrections together,

$$Vo = \frac{VP \times 273}{760 \times (273 + t)} = \frac{VP}{760\ (1 + 0.003665\ t)}$$

This formula applied to the present problem reads:

$$Vo = \frac{100 \times 727.2}{760\ (1 + 0.003665 \times 20)} = 89.2 \text{ liters.}$$

The latter calculation can be considerably simplified by using standard tables which give constants for corrections of gas volumes. These are easily obtainable and are given in part in Table IV.

According to these tables for 20° C. and 727.21 mm. Hg B.P., the factor is 0.89124; therefore:

0.89124 × 100 = 89.124 liters, 0°C. and 760 mm. Hg.
0.89124 × 4.57 = 40.7 liters of O_2 in 15 min., or 16.28 L. per hour.

The Caloric Value Calculated from the Gas Exchange.—By reference to Table V giving the heat value of 1 liter of O_2 at various respiratory quotients, it is found that at a R.Q. of 0.87, 4.888 calories are expended; 16.28 liters of O_2 is therefore equivalent to 18.4 × 4.888 = 79 calories.

The results must be calculated for surface area as well as body weight. Suppose the subject weighed 85 kg. and was 170 cm. in height; by reference to the chart for determining the surface area of man (page 540), this would be found to be 1.96 square meters. The caloric expenditure per square meter in the above case is therefore $\frac{79}{1.96} = 40.3$ calories.

TABLE I

THE PERCENTAGE OF OXYGEN WHICH IS EQUIVALENT TO THE NITROGEN FOUND IN THE EXPIRED AIR

To obtain the nitrogen in the expired air, add the percentage of CO_2 and O_2 found and subtract the sum from 100. The table gives the percentage for O_2 corresponding to this figure:

$\%N_2$	78.7	78.8	78.9	79.0	79.1	79.2	79.3	79.4	79.5	79.6	79.7	79.8
$\%O_2$	20.86	20.88	20.90	20.93	20.96	20.98	21.01	21.04	21.07	21.10	21.12	21.14
	79.9	80.0	80.1	80.2	80.3	80.4	80.5	80.6				
	21.16	21.19	21.22	21.25	21.28	21.31	21.35	21.38				

TABLE II

TENSION OF AQUEOUS VAPOR IN MILLIMETERS OF MERCURY

To obtain the dry barometer pressure, subtract the mm. Hg. corresponding to the temperature of the air from the barometer pressure at the time of the experiment:

Temp.	15°	16°	17°	18°	19°	20°	21°	22°	23°	24°	25°
Mm.	12.7	13.5	14.4	15.4	16.3	17.4	18.5	19.7	20.9	22.2	23.5

Table III

Temperature Corrections to Reduce Readings of a Mercurial Barometer with a Brass Scale to 0°C.

Subtract the appropriate quantity as found in table from the height of the barometer. The table is for a barometer with a brass scale, and the values are a little lower (about .2 mm.) than for the glass scale. The corrections for intermediate temperatures can be approximated.

Temp.	700 mm.	710 mm.	720 mm.	730 mm.	740 mm.	750 mm.	760 mm.	770 mm.
15°	1.69	1.72	1.74	1.77	1.79	1.81	1.84	1.86
20°	2.26	2.22	2.32	2.36	2.39	2.42	2.45	2.48
25°	2.83	2.87	2.91	2.95	2.99	3.03	3.07	3.11

Table IV

Table for Reducing Gaseous Volumes to Normal Temperature and Pressure

The observed volume, when multiplied by the factor corresponding to the temperature and pressure, will give the volume of the expired air reduced to 0° and 760 mm.

Mm.	15°	16°	17°	18°	19°	20°	21°	22°	23°	24°	25°
720	.898	.894	.891	.888	.885	.882	.880	.877	.873	.870	.867
730	.910	.907	.904	.901	.897	.894	.891	.888	.885	.882	.879
740	.922	.919	.916	.913	.910	.907	.904	.901	.897	.894	.891
750	.935	.932	.928	.925	.922	.919	.916	.913	.910	.907	.904
760	.947	.944	.941	.938	.934	.931	.928	.925	.922	.919	.916
770	.960	.957	.953	.950	.948	.945	.940	.936	.933	.930	.927

Table V

R. Q.	Calories for 1 liter O_2 Number	Relative calories consumed as Carbohydrate per cent	Fat per cent
0.707	4.686	0	100
0.71	4.690	1.4	98.6
0.72	4.702	4.8	95.2
0.73	4.714	8.2	91.8
0.74	4.727	11.6	88.4
0.75	4.739	15.0	85.0
0.76	4.752	18.4	81.6
0.77	4.764	21.8	78.2
0.78	4.776	25.2	74.8
0.79	4.789	28.6	71.4
0.80	4.801	32.0	68.0
0.81	4.813	35.4	64.6
0.82	4.825	38.8	61.2
0.83	4.838	42.2	57.8
0.84	4.850	45.6	54.4
0.85	4.863	49.0	51.0
0.86	4.875	52.4	47.6
0.87	4.887	55.8	44.2
0.88	4.900	59.2	40.8
0.89	4.912	62.6	37.4
0.90	4.924	66.0	34.0
0.91	4.936	69.4	30.6
0.92	4.948	72.8	27.2
0.93	4.960	76.2	23.8
0.94	4.973	79.6	20.4
0.95	4.985	83.0	17.0
0.96	4.997	86.4	13.6
0.97	5.010	89.8	10.2
0.98	5.022	93.2	6.8
0.99	5.034	96.6	3.4
1.00	5.047	100.0	0.0

(From Lusk.)

CHAPTER LXIII

STARVATION

In order to furnish us with a standard with which we may compare other conditions, we shall first of all study the metabolism during starvation. A valuable chart compiled from observations made in the Carnegie Institution of Washington on a man who fasted for thirty-one days is reproduced in Fig. 182.

The Excretion of Nitrogen.—When an animal is starved, it has to live on its own tissues, but in doing so it saves its protein, so that the excretion of nitrogen falls after a few days to a low level, the energy requirements being meanwhile supplied, so far as possible, from stored carbohydrate and fat. Although always small in comparison with fat, the stores of carbohydrate vary considerably in different animals. They are much larger in man and the herbivora than in the carnivora. *During the first few days of starvation* it is common, in the herbivora, to find that the excretion of nitrogen is actually greater than it was before starvation, because the custom has become established in the metabolism of these animals of using carbohydrates as the main fuel material, so that when carbohydrates are withheld, as in starvation, proteins are used more than before and the nitrogen excretion becomes greater. We may say that the herbivorous animal has become carnivorous. The same thing may occur in man when the previous diet was largely carbohydrate; thus, almost invariably in man the nitrogen output is larger on the third and fourth days of starvation than on the first and second.

Another factor influencing the nitrogen excretion during the early days of the fast is the amount of previous intake of nitrogen; the greater this has been, the greater the excretion. By the seventh day, however, a uniform output of nitrogen will usually be reached irrespective of the individual's protein intake. During the greater part of starvation, most of the energy required to maintain life is derived from fat, as little as possible being derived from protein. This type of metabolism lasts until all the available resources of fat have become exhausted, when a more extensive metabolism of protein sets in, with the consequence that the nitrogen excretion rises. This is really the harbinger of death—it is often called the *premortal rise in nitrogen excretion.* It indicates that all the ordinary fuel of the animal economy has been used up, and that it has

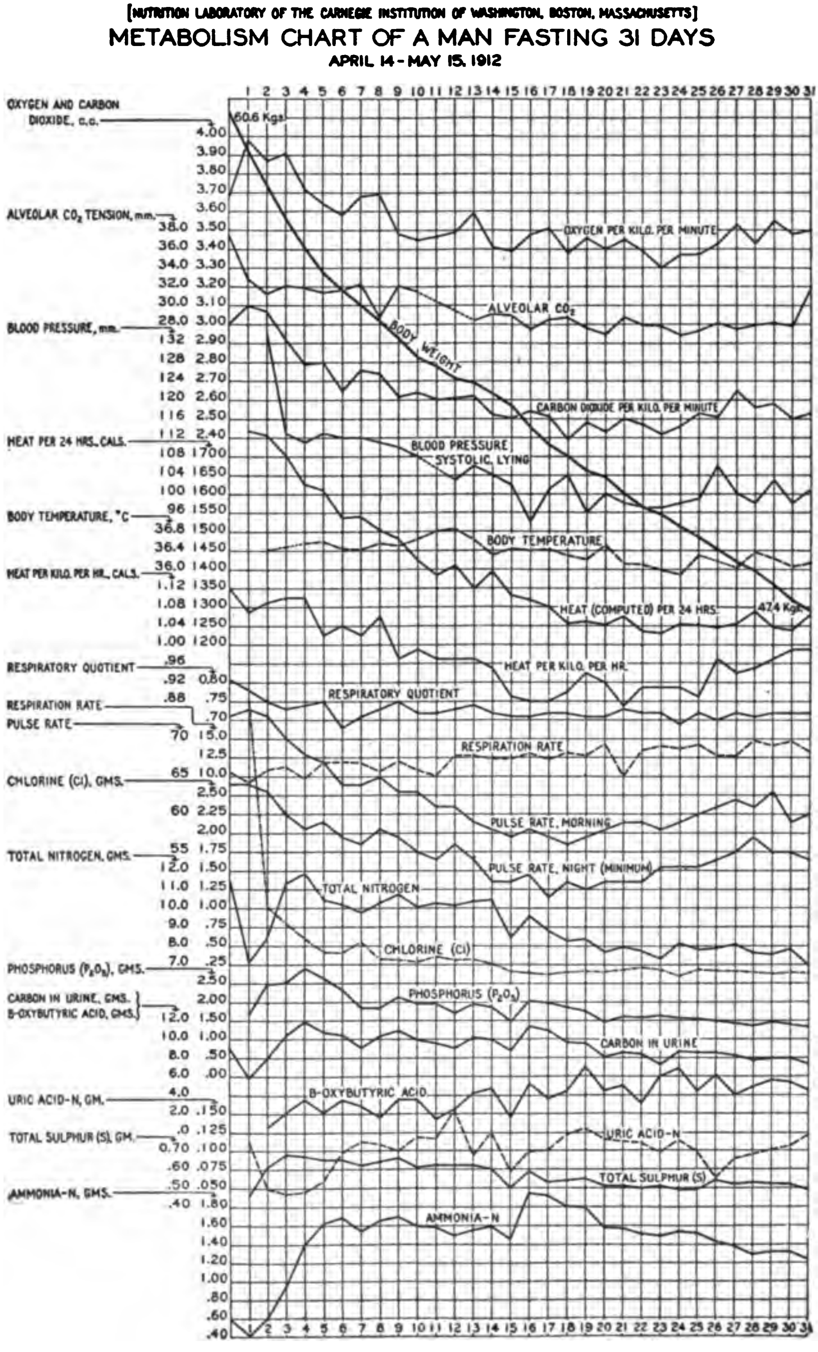


Fig. 182.—Curve constructed from data obtained from a man who fasted for thirty-one days. The days of the fast are given along the abscissae, and the various measurements along the ordinates. (From F. G. Benedict.)

become necessary to burn the very tissues themselves in order to obtain sufficient energy to maintain life. Working capital being all exhausted, an attempt is made to keep things going for a little longer time by liquidation of permanent assets. But these assets, as represented by protein, are of little real value in yielding the desired energy because, as we have seen, only 4.1 calories are available against 9.3, obtainable from fats.

These facts explain why during starvation a fat man excretes daily less nitrogen than a lean man, and why the fat man can stand the starvation for a longer time. The premortal rise is, however, not prevented by feeding oil, which would seem to indicate that death may be due not so much to the absence of fuel as to serious nutritional disturbance of essential organs; e. g., there may be no available material to supply the glands of internal secretion with the building stones they must have (see page 580).

Not only is there this general saving of protein during starvation, but there is also a discriminate utilization of what has to be used by the different organs, according to their relative activities. This is very clearly shown by comparison of *the loss of weight which each organ undergoes during starvation.* The heart and brain, which must be active if life is to be maintained, lose only about 3 per cent of their original weight, whereas the voluntary muscles, the liver and the spleen lose 31, 54 and 67 per cent, respectively. No doubt some of this loss is to be accounted for as due to the disappearance of fat, but a sufficient remainder represents protein to make it plain that there must have been a mobilization of this substance from tissues where it was not absolutely necessary, such as the liver and voluntary muscles, to organs, such as the heart, in which energy transformation is *sine qua non* of life. The vital organs live at the expense of those whose functions are accessory.

The **energy output** per square meter of body surface steadily declines. In the man examined by Benedict, it was 958 C. per square meter of surface at the end of the first twenty-four hours, but only 737 on the thirty-first day of the starvation period. The oxygen intake and carbon-dioxide output correspondingly diminish.

The behavior of the **nitrogenous metabolites** in the urine is of particular interest, the following facts being of significance: Urea nitrogen relatively falls and NH_3-N rises. For example, on the last day of feeding the percentage output of NH_3-N in relation to total nitrogen was 3.16; on the eighth day of the fast it was 14.88 (Cathcart).[2] Acidosis is the cause. The total amount of creatinine and creatine shows only a slight fall, but creatinine relatively decreases and creatine increases (Cathcart). Since creatine is a substance peculiar to muscle tissue, it is possible by

comparing the creatine and creatinine output with that of nitrogen to determine whether all of the nitrogen liberated by the breakdown of muscle has been excreted, or whether some has been retained either for resynthesis in the muscle itself or for use elsewhere. If the muscle breakdown as calculated from the creatine-creatinine output is greater than that calculated from the nitrogen, synthesis of the noncreatine remainder must be occurring; whereas if the breakdown calculated from nitrogen is greater than that calculated from creatine, etc., other tissues than muscle must be contributory. Stored nitrogen or free nitrogen in transit from tissue to tissue for utilization is the most likely source of such excess nitrogen.

That transference of nitrogenous substances from place to place in the body in starvation is proved (1) by the constant presence of amino nitrogen in the blood and tissues (Van Slyke); and (2) by the effect of copious water drinking. The latter causes a decided increase in the output of nitrogen, but it does not appear that the extra nitrogen is due to increased protein breakdown. It is probable, however, that in such cases there would also be an increase in endogenous protein metabolism, since the washed-out free nitrogen would have to be replaced.

Excretion of Purines.—Although at first they fall somewhat, the total amount increases as the fast progresses. Perhaps the first decline is due to general using up of hypoxanthine of muscle and the later rise to the breakdown of nuclei (page 638).

Excretion of Sulphur.—It is important to compare the excretion of sulphur and nitrogen. In the early days of starvation a ratio of 17 N: 1 S has been found, but later one of 14.5: 1, which is practically the same as that in muscle (i. e., 14; 1), indicating that late in fasting the main source of protein supply is muscle.

Several of the changes observed during starvation can be attributed to the condition of *acidosis* which supervenes. The acids are derived from incomplete combustion of fat (see page 683), and are represented by β-oxybutyric, the amount being sometimes considerable (10-15 grams a day), especially in obese individuals. The large ammonia excretion (sometimes 2 grams a day) is evidently for the purpose of neutralizing the excess of acid. Another consequence of the acidosis is the decline in the alveolar tension of CO_2 (page 354), and it is possible that some of the circulatory changes shown in the chart may also be dependent on it. The method of repeated fasting used for reducing obesity is quite safe if the acidosis is carefully watched.

Many *secondary changes* also occur in the starving organism. Thus, the mobilization of fat is often responsible for a pronounced increase in the fat content of the blood (see page 698), and that of protein explains

the presence of an amount of amino nitrogen not much below that of normal animals (viz., 4 mg. per 100 c.c. of blood). Similarly with carbohydrates, early in the condition the blood sugar becomes much lower than normal, but then remains steady. This is significant when we remember that after two or three days of starvation all of the available glycogen has been used up. It indicates that carbohydrate must be essential for life, and that it is produced in starvation from proteins (see page 667).

Starvation ends in death in an adult man in somewhat over four weeks but much sooner in children, because of their more active metabolism. At the time of death the body weight may be reduced by 50 per cent. The body temperature does not change until within a few days of death, when it begins to fall, and it is undoubtedly true that if means are taken to prevent cooling of the animal at this stage, life will be prolonged.

Death from starvation must be due either to a general failure of all the cells or to injury of certain organs that are essential for life. Since the loss of protein from the body as a whole may vary between 20 and 50 per cent at the time of death by starvation, it is unlikely that general failure can be its cause. If it were so, death would always occur when some fixed loss of protein had occurred. Certain organs evidently cease to perform their function, either because they are deprived of raw material for the elaboration of some substance (hormone) necessary for life, or because the organs themselves wear out from want of nourishment.

NORMAL METABOLISM

Apart from the practical importance of knowing something about the behavior of an animal during starvation, such knowledge is of great value in furnishing a standard with which to compare *the metabolism of animals under normal conditions.* Taking again the nitrogen balance as indicating the extent of protein wear and tear in the body, let us consider first of all the conditions under which equilibrium may be regained. It would be quite natural to suppose that, if an amount of protein containing the same amount of nitrogen as is excreted during starvation were given to a starving animal, the intake and output of nitrogen would balance. We are led to make this assumption because we know that any business balance sheet showing an excess of expenditure over income could be met by such an adjustment. But it is a very different matter with the *nitrogen balance sheet* of the body; for, if we give the starving animal just enough protein to cover the nitrogen loss, we shall cause the excretion to rise to a total which is practically equal to

the starvation amount *plus* all that we have given as food; and although by daily giving this amount of protein there may be a slight decline in the excretion, it will never come near to being the same as that of the intake. The only effect of such feeding will be to prolong life for a few days.

Nitrogenous Equilibrium.—To attain equilibrium we must give an amount of protein whose nitrogen content is at least two and one-half times that of the starvation level. For a few days following the establishment of this pure protein diet, the nitrogen excretion will be far in excess of the intake, but it will gradually decline until the two practically correspond. Having once gained an equilibrium, we may raise its level by gradually increasing the protein intake. During this progressive raising of the ingested protein, it will be found, at least in the carnivora (cat and dog), that a certain amount of nitrogen is retained by the body for a day or so immediately following each increase in protein intake. The excretion of nitrogen, in other words, does not immediately follow the dietetic increase. The amount of nitrogen thus retained is too great to be accounted as a retention of disintegration products of protein; it must therefore be due to an actual building up of new protein tissue—that is, growth of muscles.

Nitrogenous equilibrium on a protein diet alone is readily attainable in the cat, and less readily in the dog. But in man and the herbivorous animals, it is impossible to give a sufficiency of protein alone to maintain equilibrium; there will always be an excess of excretion over intake. Indeed it scarcely requires any experiment to prove this, for it is self-evident when we consider that there are less than 1000 C in a pound of uncooked lean meat, and that there are few who could eat over three pounds a day, an amount, however, which would scarcely furnish all of the required calories. A person fed exclusively on flesh is therefore being partly starved, even although he may think that he is eating abundantly and be quite comfortable and active. This fact has a practical application in the so-called *Banting cure for obesity.*

Protein Sparers.—Very different results are obtained when carbohydrates or fats are freely given with the protein to the starving animal. Nitrogen equilibrium can then be regained on very much less protein, so that we speak of fats and carbohydrates as being *"protein sparers."* Carbohydrates are much better protein sparers than fats; indeed they are so efficient in this regard that it is now commonly believed that carbohydrates are essential for life, and that when the food contains no trace of carbohydrates, a part of the carbon of protein has to be converted into carbohydrate. This important truth is supported by evidence derived from other fields of investigation (e. g., the behavior of

diabetic patients, in whom the power to use carbohydrates is greatly depressed). The marked protein-sparing action of carbohydrates is illustrated in another way—namely, by the fact that we can greatly diminish the protein breakdown during starvation by giving carbohydrates. In this way we can indeed reduce the daily nitrogen excretion to about one-third its amount in complete starvation. Carbohydrate starvation is said to entail a failure of the muscles to use again in their metabolism certain of the products (e. g., creatine) which result from their disintegration. At any rate it has been found that creatine is excreted in the urine when no carbohydrates are available.

In the case of man living on an average diet, although the daily nitrogen excretion is about 15 grams, it can be lowered to about 6 grams provided that in place of the protein that has been removed from the diet enough carbohydrate is given to bring the total calories up to the normal daily requirement. If an excess of carbohydrate over the energy requirements is given, the protein may be still further reduced without disturbing the equilibrium. It has been found that it is not the amount of carbohydrate alone that determines the ease with which the *irreducible protein minimum* can be reached; the kind of protein itself makes a very great difference. This has been very clearly shown by one investigator, who first of all determined his nitrogen excretion while living exclusively on starch and sugar, and who then proceeded to see how little of different kinds of protein he had to take in order to bring himself into nitrogenous equilibrium. He found that he had to take the following amounts: 30 gm. meat protein, 31 gm. milk protein, 34 gm. rice protein, 38 gm. potato protein, 54 gm. bean protein, 76 gm. bread protein, and 102 gm. Indian-corn protein. The organism is evidently able to satisfy its protein demands much more readily with meat than with vegetable proteins.

This variability in the food value of different proteins depends on their ultimate structure—that is, on the proportion and manner of linkage of the various amino acids that go to build up the molecule. In no two proteins are these building stones, as they are called, present in exactly the same proportions, some proteins having a preponderance of one or more and an absence of others, just as in a row of houses there may be no two that are exactly alike, although for all of them the same building materials were available. Albumin and globulin are the most important proteins of blood and tissues, so that the food must contain the necessary units for their construction. If it fails in this regard, even to the extent of lacking only one of the units, the organism will either be unable to construct that protein, and will therefore suffer from partial starvation, or it will have to construct for itself this missing unit. It

growth—vitamines?), to which were added more carbohydrate, purified fat, and the protein whose influence on growth it was desired to study. The same diet was fed at regular intervals to a given batch of rats, and the weight of each rat was periodically taken, the observation being prolonged until the animals grew to maturity and produced young, and these again grew to maturity, reproduced, and so on. By plotting the results in curves, with the time periods along the abscissæ and the average weight of the rats of each batch along the ordinates, the extent of the influence of a given diet on *the curve of growth* was obtained. A normal curve of growth is shown in No. 1 of Fig. 183. It was obtained from results secured by adding liberal amounts of casein to the basal diet.

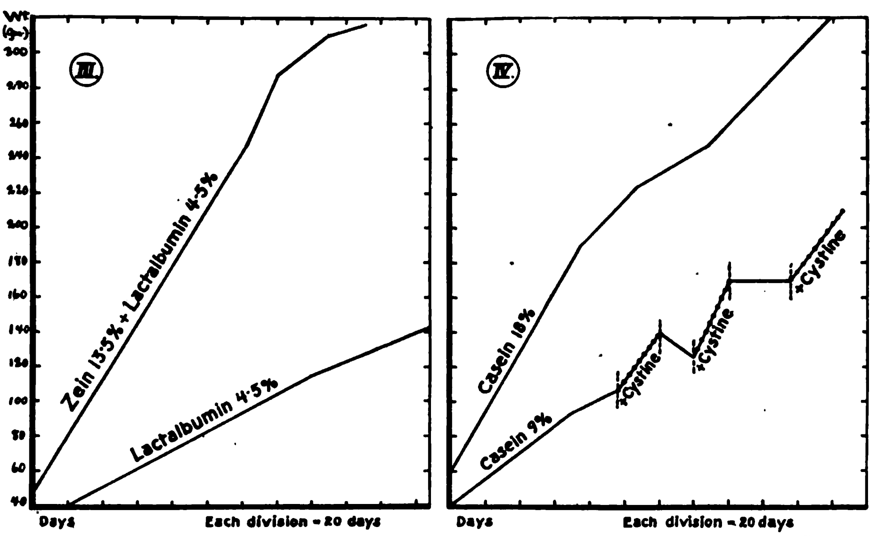


Fig. 184.—Curves of growth of rats on basal rations plus the proteins indicated. In curve III the effect of the addition of zein to an inadequate allowance of the perfect protein, lactalbumin, is shown; and in IV the effect of the addition of cystine to a deficient casein allowance. (From Lafayette B. Mendel and T. B. Osborne.)

Similar curves were obtained with lactalbumin of milk and ovalbumin and ovovitellin of egg. Perhaps the most interesting substances capable of producing the normal curve of growth are certain of the proteins that T. B. Osborne has succeeded in separating in crystalline form from vegetable foodstuffs. These are edestin (hempseed), globulin (squash seed), excelsin (Brazil nut), glutelin (maize), globulin (cottonseed), glutein (wheat), glycinin (soy bean), cannabin (hempseed).

That growth proceeds normally with any one of these proteins when fed abundantly does not, however, necessarily indicate that each contains in adequate proportion all of the necessary units to meet the protein demands of growing tissues. In the case of casein, for example, one of the units, namely, glycocoll, which is the simplest of all the

amino acids, is entirely missing, and another, cystine, which is a sulphur-containing amino acid, is present only in small amount. The absence of glycocoll, however, is not of importance, because the organism can manufacture it for itself (see page 630). In the case of cystine, which the tissues can not manufacture themselves, the deficiency has to be made up for by feeding an excess of casein so as to cover the needs of the tissues for this amino acid. By so doing a superabundance of most of the other units will be ingested, and this superabundance will entail the destruction and excretion of the useless amino acids, a process, however, which is conducted in such a way as to permit of the utilization, by the organism, of a part of the energy which the cast-off amino acids contain (see page 667). It is, therefore, not entirely a wasteful process.

When the supply of casein is limited, on the other hand, the curve of growth becomes subnormal, because an insufficient supply of cystine is thereby offered (Fig. 184). Similar results have been obtained in the case of edestin, a protein from hempseed. This contains an insufficiency of the diamino acid, lysine. Fed in abundance, edestin gave a normal curve of growth, but when fed in insufficient amount the curve failed to ascend properly, which, however, it could be made to do by adding some lysine to the edestin.

There is a large group of proteins which fail to permit of any growth no matter in what amounts they may be added to the basal ration. These include: legumelin (soy bean), vignin (vetch), gliadin (wheat or rye), legumin (pea), legumin (vetch), hordein (barley), conglutin (lupine), gelatine (horn), zein (maize), phaseolin (kidney bean). The adequacy to maintain growth of any of these pure proteins varies according to the deficiency in their amino acids. In the case of gliadin of wheat or rye, glycocoll is lacking, and lysine is present only in small amount (see table). The absence of glycocoll can not, however, as we have already seen in the case of casein, explain the inadequacy of gliadin as a foodstuff for growth (Curve II in Fig. 183). It must be the lysine that is at fault. A still more deficient protein is the *zein* of maize. With this as the only protein added to the basal diet, the curve of growth actually descends (Curve III of Fig. 183), thus indicating that the animal is starving and must soon succumb. The missing units in this protein are glycocoll, lysine and tryptophane (see table on page 575), and it is very significant that if the latter two amino acids are supplied along with zein, an almost normal curve of growth will result. Some improvement can even be brought about by giving tryptophane alone; that is to say, the curve assumes a horizontal line instead of descending, indicating that, although inadequate for growth, the diet is now sufficient for the maintenance of life.

The important fact demonstrated by these experiments, is that *certain diets are adequate for the maintenance of life although they are inadequate for growth.* In conformity with this conclusion, it was found when young white rats were fed with gliadin alone for periods of time ex-

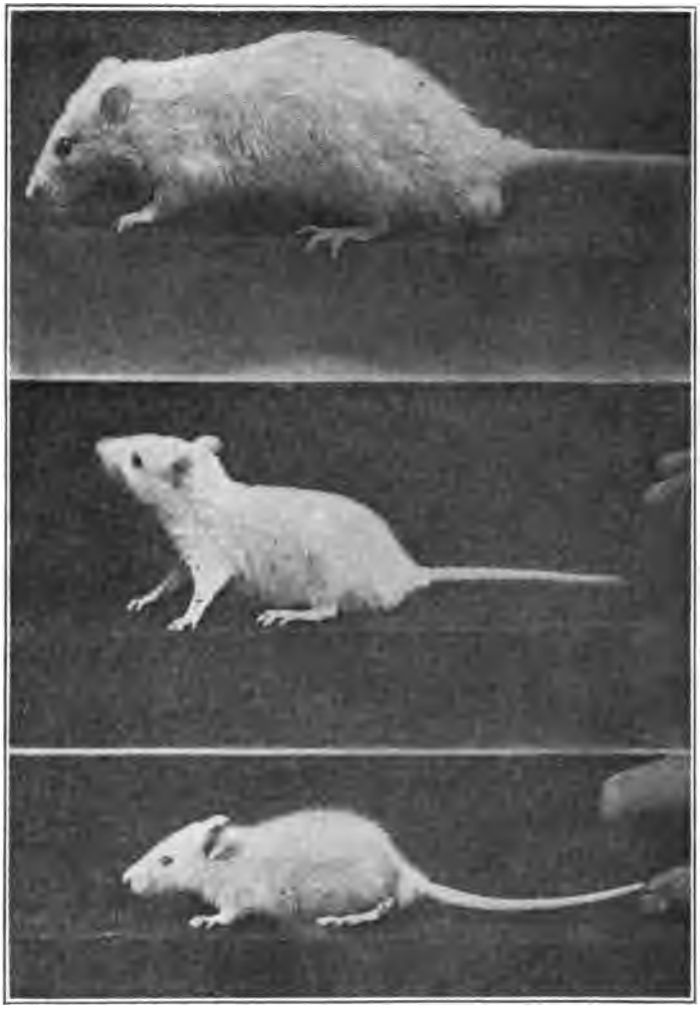


Fig. 185.—Photographs of rats of same brood on perfect diet (uppermost picture); on a maintenance diet but inadequate for growth (middle picture); and on a diet that was inadequate both for maintenance and growth. (From Mendel and Osborne.)

ceeding those in which they should have become full grown, that they remained in an ungrown stunted condition. The capacity to grow had not, however, been lost, for when the gliadin was replaced by milk, the animals resumed growth at a very great rate. The capacity to grow

had only been *inhibited* by the inadequate diet, and there was nothing really abnormal about the stunted animals. For example, the reproductive function developed normally, as was shown in the case of a young female rat which, after being fed with gliadin as the sole protein supply for 154 days, was mated and produced four young. Although the mother was still maintained on the gliadin diet, the young rats presented normal growth, for they were living on the milk supplied by the mother, and this milk, because it contained either casein or some other necessary accessory factor (vide infra), was an adequate food.

After removal from the mother, three of these rats were fed on an artificial diet of casein, edestin and the basal ration, and continued the normal course of growth, but when one of them was placed on the gliadin food mixture it immediately failed to grow properly. It would appear from these experiments that, of the two amino acids that are missing or deficient in gliadin—namely, glycocoll and lysine—it must be the lysine that is essential for growth. This very important conclusion was fully corroborated by finding that, in young rats stunted by previous gliadin feeding, growth immediately started when lysine was added to the diet and ceased again when the lysine was removed, and so on, the experiments being often repeated in various modifications. Mendel and Osborne call attention to the relatively high percentage of lysine in all those proteins that are concerned in nature with the growth of young animals; thus, it is present in large amounts in casein, lactalbumin and egg vitellin.

It is particularly in protein of vegetable origin that indispensable units are likely to be missing, the best known of these units being the aromatic amino acids, tyrosine and tryptophane; the diamino acid, lysine; and the sulphur-containing acid, cystine. Some animal proteins, such as gelatine, also fail to contain aromatic groups, and are therefore utterly inadequate as protein foods.

That the absence of one or two units should render a protein utterly incapable of maintaining life suggests that a specific role may be taken by certain amino acids in the maintenance of nutritional rhythm; thus, they may be necessary for the elaboration of some hormone or other internal secretion essential to life, such as epinephrine, the active principle of the suprarenal gland. This is an aromatic substance not far removed in its chemical structure from tyrosine (see page 734). It is therefore natural to suppose that the absence of the tryptophane unit in zein is the reason that this protein is incapable of maintaining the initial body weight.

In attacking the problem from this viewpoint, Hopkins and Willcock[10] made observations on the *survival period of young mice;* that is, the period during which the animals survived when fed on a basal diet

mixed either with zein alone or with zein *plus* small quantities of tryptophane. It was found that, with zein alone, the mice were unable to maintain growth; they lost in weight and died in from about a week to about a month. Other mice fed on the same amount of basal diet and zein, but to which was also added some tryptophane, although they did not grow, were capable of maintaining their body weight and lived in some instances for nearly a month and a half. There were other indications of the difference in the efficiency of the two diets. The mice fed on the zein alone were very inactive, and remained for a considerable period of the time in a condition of torpor. The hair was ruffled, the eyes were half closed, and the ears, feet and tail were cold. The animals, however, gave evidence of having a good appetite. On the other hand, the mice to which tryptophane was also given manifested a strikingly different behavior, being active and more or less normal until just before death. That both groups of animals failed to live more than forty-four or forty-eight days is probably to be accounted for by the absence in the zein of the other unit, lysine. Had this been added along with the tryptophane it is probable, in the light of Mendel and Osborne's observations, that the animals would have survived much longer.

To supply the missing unit, besides using the pure amino acid, we *may employ other proteins which contain the required amino acid* (Curve III of Fig. 184). That mixtures of protein foodstuffs are desirable has long been apparent to those who have studied practical dietetics. We must combine the unsuitable protein with others which, although in themselves perhaps also unsuitable, yet furnish us with a mixture which contains all the essential units both for maintenance and growth. As Mendel points out, these considerations suggest that we may be able to utilize certain of the low priced protein by-products of the cereal, meat and milk industries. The test of the adequacy of the corrected diet must, however, be determined by experiments of the type which we have just described. It is probably in stock-raising rather than in connection with human nutrition that these facts will prove of practical value; for, not only is the diet of man more varied, but it contains animal proteins in which the deficiencies are not so common.

Most important work of this character is being conducted by McCollum and his collaborators.[12] It would take us beyond the confines of this book to discuss the results in detail, but it may be mentioned that they have shown that, since the adequacy of the diet depends on a multiplicity of factors besides the amino-acid make-up of proteins,—some of which we shall discuss immediately,—very extensive observations with various food mixtures must be conducted over long periods of time. The nutritive values of the common cereals added to a standard diet that had brought the animals (rats) to the threshold of death,

were found to be as follows: With cornmeal there was immediate recovery and rapid growth, both of which were also secured in considerable degree by wheat embryo and entire wheat kernel; with rye and oats, on the other hand, there was little if any improvement.

Much work is, of course, yet to be done before we can determine the exact role which each unit plays in the physiologic development of young animals. *To sum up* what we already know, it may be said that glycocoll is not essential, since it can be manufactured by the animal itself; that tryptophane is essential for maintenance, probably because it is required for the production of certain essential hormones, for the make-up of which in its absence other tissues must become disintegrated, leading therefore to a diminution in body weight; and that lysine appears to be essential for growth. Tissues can be maintained without lysine, but they can not grow, for the slight trace which most food contains of this important amino acid may be sufficient for maintenance purposes, but utterly inadequate for growth. That the young rats in the experiments of Mendel grew normally while living on milk supplied by the stunted mother indicates that the requisite lysine must have been produced in the mother's body.

In the application of the foregoing principles to human dietetics, it is undoubtedly safe to follow Bayliss's advice to take care of the calories and allow the proteins to take care of themselves.[11] For example, in the case of milk the deficiency of cystine in its chief protein, casein, is corrected by the presence of lactalbumin, which, though present in only small amounts, contains sufficient quantities of this amino acid to meet the demands of the growing tissue.

These observations on maintenance and growth suggest very interesting applications in connection with *the growth of tumors.* Is it possible that we might retard the growth of tumors by a diet that was insufficient for growth while sufficient for maintenance. In an experiment devised to test this proposition mice were fed on a diet of starch, lard, lactose and gluten on which they could merely maintain existence but failed to grow. Some of these rats were inoculated with a rapidly growing tumor at the same time as another batch of mice kept on normal diet, and it was found that the tumor grew much more slowly in the stunted mice than in the others. One mouse, for example, on the restricted diet had a scarcely visible tumor 52 days after the inoculation. When this mouse, however, was placed on a normal diet of bread, milk, etc., the tumor immediately began to grow at a very great rate.[13] Too much importance should not be placed on this experiment.

We shall now pass on to consider some of the factors besides the protein content which have an important bearing on dietetic efficiency.

CHAPTER LXV

NUTRITION AND GROWTH (Cont'd)

THE RELATIONSHIP OF OTHER FACTORS THAN PROTEINS

The Relationship of Carbohydrates.—As we have seen elsewhere, carbohydrates are almost certainly essential for normal metabolism. If they are not given with the food, they must be manufactured out of protein by the organism itself. It is not surprising, therefore, that their absence from the diet of growing animals should lead to abnormality in the rate of growth. Pediatrists have not infrequently insisted that one form of carbohydrate is more advantageous for growth than another. This no doubt in the main is true, but the whole question of adequacy probably depends on the digestibility of the carbohydrate and not upon its essential chemical nature. It is likely that the only carbohydrate required by the tissues is glucose. The readiness with which the carbohydrate of the food becomes converted into this monosaccharide is probably the only determinant of its efficiency as food material.

The Relationship of Fats.—Although fats are an invariable constituent of practically every diet, it is yet a debatable question as to whether they are essential to the maintenance of a healthy normal organism. Difficulties standing in the way of a solution of this problem are that it is not only technically very difficult to remove fat entirely from the common foodstuffs, but also that the simple fats are usually associated with substances having similar solubilities and physical properties: namely, the lipoids, phosphatides, cholesterol, pigments, etc. Since these substances are present in practically every cell, it is almost certain that they can be manufactured by living protoplasm. Indeed, experimental evidence is not wanting to show that this is actually the case. Although the cell can manufacture lipoids, a young animal can apparently not grow when these substances, as well as simple fat, are entirely absent from the diet. This has been shown by feeding young mice on a diet from which all traces of fat and lipoids had been removed by extraction with alcohol and ether (Stepp)[14]. On such a diet the mice lived only a few weeks. They could be kept alive much longer when some of the alcohol-ether extract was mixed with the diet, but not so when neutral fat instead of the alcohol-ether extract was added. The

addition of the ash of the lipoid extract failed to maintain the mice, so that the lacking substance could not be inorganic in nature.

More recent and extended observations, however, have shown that neutral fat is also necessary for the adequate and continued growth of the animal. For a period of two months or so an animal may, as we have seen from Osborne and Mendel's experiments, grow in apparently normal fashion on an artificial fat- and lipoid-free diet composed of casein, carbohydrate and inorganic salts, but sooner or later the great majority of these animals begin to show failure of adequate growth. The inadequately growing animals often manifest indications of malnutrition other than the failure to increase in weight; for example, inflammation of eyes, roughening of the fur, etc. When certain fats are added to the inadequate diet, normal growth is immediately resumed. Fats producing this normal growth are such as butter fat, or the fat extracted from egg yolk, or cod-liver oil, added to the extent of 5 per cent of the ration. On the other hand, vegetable oils, such as olive oil or almond oil, are inefficient in promoting growth. That all oils or fats do not suffice to produce growth, and that one dose of an adequate oil or fat may be sufficient to stimulate it, indicate that something other than the mere presence of the comparatively simple fat molecule—that is, some accessory material—must be the agency responsible for the growth.

This conclusion is further supported by the interesting observation of McCollum and Davis that vegetable oils can be rendered efficient for growth by shaking them with a solution of soap prepared by completely saponifying butter fat with potassium hydroxide in the absence of water.

ACCESSORY FOOD FACTORS, VITAMINES

In searching for the nature of the accessory food factors, the important observations which have been made in recent years concerning the so-called vitamines must be considered. These are substances essential in the diet for the proper maintenance of nutrition in adult animals.

The existence of such substances was suggested by observations on the disease *beriberi*, which is caused by exclusive feeding on polished rice; that is, on rice from which the pericarp had been removed by the process of polishing. When patients suffering from this disease were given unpolished rice, the symptoms immediately disappeared. Further investigation of the exact nature of these substances was greatly facilitated by the discovery that a similar condition is readily induced by feeding fowls on polished rice. The birds develop a polyneuritis, from which, however, they very promptly recover if some rice polishings or,

better still, an extract of rice polishings, is added to the polished rice diet. The extract is made by means of slightly acid 91 per cent alcohol, and from it Funk has succeeded in separating a substance in crystalline form apparently related to the pyrimidines, which it will be remembered are a characteristic constituent of the nucleins. Doses as small as 0.02 to 0.04 gm. of this material given by mouth were adequate to cure the polyneuritis of fowls in from six to twelve hours; indeed, in some cases the bird seemed quite well after three hours. A similar substance has also been extracted from yeast, milk, brain and lime juice, and it has been called, for want of a better name, vitamine.[15]

It is quite likely that other diseases, such as scurvy, may also be due to the absence of some vitamine in the diet—some substance, namely, which in the case of this particular disease would seem to be absent in preserved food, the continued taking of which is so frequently its cause. Fresh fruit and other foods added even in small amounts to such a diet would appear to supply the necessary vitamine.

It is not the higher animals alone that suffer from the want of some such substance as vitamine. It has been shown, for example, that, when a normal artificial culture medium is inoculated with yeast in very small amounts, it fails to grow, whereas the same quantity will grow luxuriantly in a medium to which *sterilized* beer wort has been added. Vitamine is not of the nature of a ferment, since it withstands heating to 120° C. for more than an hour. The addition of yeast to dietaries that are deficient in vitamines is an excellent corrective.

Returning now to the accessory substances that seem to be adherent to certain forms of fat, we see at once that they can not be exactly the same as the so-called vitamine of Funk, for they contain no nitrogen. There are, therefore, probably two accessory factors concerned in adequate growth. One of these must be present in the protein-free milk which serves as a constituent of the basal diet used in Osborne and Mendel's experiments, for we have seen that animals will grow on this for a certain period, provided the proper amino acids are present. Later, however, they pass into a state where there is no growth but adequate maintenance. If now the other accessory factor is added, as, for example, butter fat or a small amount of milk itself (i. e., in place of protein-free milk), then growth will be resumed at its normal rate. "Either of the determinants may become curative. Both are essential for growth when the body store of them becomes depleted." McCollum suggests that these accessory factors should at present be called the "fat-soluble A" and "water-soluble B." The latter is present in yeast cells, in fat-free milk, and in many other animal foods, and is probably the same as Funk's vitamine. The former is soluble in the fat solvents,

being present in most animal fats, but not in all; for example, it is absent from the fat surrounding the pig's heart. By using such a nomenclature it is recognized that the subject is as yet only in an early state of development.

We may sum up the main facts of this chapter by stating that *growth and maintenance are more than a mere problem of energy supply. Granted that this is sufficient, we must alsa have a suitable admixture of building units of protein and the presence of extremely small quantities of some unknown accessory substances.* These are present in some natural foods but not in others, and some are soluble in water and others in fats. They are found, for example, in animal fats but not in those of vegetable origin. Both fat- and water-soluble factors are present in large quantities in milk.

Both accessory food factors are necessary, as is illustrated in the following summary of experiments from Lusk's "Science of Nutrition," (third edition).

Purified	protein	+	carbohydrate	+ vegetable fat	+	inorganic	salts	= no growth.
"	"	+	"	+ butter fat	+	"	"	= no growth.
"	"	+	"	+ vegetable fat	+	"	"	+ vitamines (accessory factor B) = no growth.
"	"	+	"	+ butter fat	+	"	"	+ vitamines = growth.

The Relationship of Inorganic Salts.—Inorganic salts are also an essential ingredient of the diet. McCollum found that young animals soon ceased to grow when fed on a diet of corn and purified casein, but that rapid growth returned when a suitable salt mixture was added. Oats, wheat, and beans have also been shown to require some adjustment of their ash content to make them adequate for growth. Most of the animal foods contain in themselves sufficient inorganic material, as is evidenced among other things by the adequacy of milk alone as diet for growing animals and the abhorrence of salt that is shown by strictly carnivorous animals. In the usual mixed diet of man there is almost always enough inorganic material, the salt which he adds being largely for seasoning purposes. When a preponderance of vegetable food is taken, however, the salt comes to have a real dietetic value.

The practical application of the results of these numerous and at present somewhat bewildering observations to the nutrition of man, and particularly to the dietetic treatment of disease, is undoubtedly very great. This is especially so in infants and growing children, in whom the correction of some slight inadequacy in the diet may have the most pronounced results, not only on growth and nourishment, but also on the power of resistance against disease and infection. The beneficial influence of cod-liver oil, for example, may depend on some fat-

soluble accessory food factors, while the miraculous benefit which scorbutic children derive from the addition of the juice of limes, lemons, etc., to the food is undoubtedly due to such influences. The accumulating mass of evidence as to the faulty nutrition in animals fed on single kinds of food that fail to contain both kinds of food factors emphasizes the necessity in the dietetic treatment of such diseases as diabetes, nephritis, etc., of seeing to it that the diet is sound, not only in calories, protein content, and palatability, but also with regard to the presence of accessory food factors.

CHAPTER LXVI

DIETETICS

THE CALORIE REQUIREMENT

In the application of the important facts that have been reviewed in the preceding chapters to the science of *dietetics,* the question arises as to how we may determine with scientific accuracy just exactly *how much food should be taken under varying conditions of bodily activity.* In a general way, we know that the amount of food that we require to take is proportional to the nature and amount of bodily exercise that is being performed at the time; and that, if the food supply is inadequate, the work before long will fall off not only in quantity but in quality as well. "Horses (also men) work best when they are well fed, and feed best when they are well worked," is an old adage and one the truth of which can not be overestimated in the consideration of all questions of dietary requirements. An ill-fed beggar will rather suffer from the pain and misery of starvation than attempt to perform a piece of work that the well-meaning housewife bargains should be done before she gives him a meal. The spirit may be willing but the flesh is weak. If he could be trusted, he should be fed first and worked afterwards. Besides the amount of work, two other factors are well known to influence the demand for food—namely, growth and climate. A young, growing boy will often demand as much if not more food than would appear to be his proper share, from a comparison of his body weight with that of his seniors; and, other things being equal, it is well known that we are inclined to eat much more heartily of food during the cold days of winter than during the sultry days of July and August.

That we know these facts in a general way, indicates that the first step to take in the exact determination of dietetic requirements is to find out how much energy the body expends under varying conditions of activity. This, as we have seen, may be done by having the person live for some time in a respiration calorimeter, so that we may measure the calorie output. To the conclusions drawn from results of observations made under such artificial and unusual conditions of living, the objection might, however, quite justly be raised that they need not apply to persons going about their ordinary routine of life. To meet

this objection another method, which we may call the *statistical*, is available. It consists in taking the average diet of a large number of individuals and comparing the calorie value with the average amount and type of work that they are meanwhile called upon to perform, and can best be used where the diet is accurately known, as in public institutions, the army, the navy, etc. The total food supplied is then divided by the number of individuals, this giving the per capita consumption. Obviously some get more than others, but when a sufficient number of individuals is included, such errors become eliminated by the law of averages.

The reliability of this method is testified to by the remarkable correspondence in the calorie value of the food consumed by farmers in widely different communities:

	Calories
Farmers in Connecticut	3,410
" " Vermont	3,635
" " New York	3,785
" " Italy	3,565
" " Finland	3,474
Average	3,551*

*Lusk: The Fundamental Basis of Nutrition.

The average inhabitant of various cities:

London	2,665
Paris	2,903
Munich	3,014
Königsberg	2,394**

**Rubner.

Individuals in different callings:

Farmers' families (U.S.A.)	3.560
Mechanics' families (U.S.A.)	3,605
Professional men's families (U.S.A.)	3,530
Army (U.S.A.)	3,851
Navy (U.S.A.)	4,998†

†Atwater.

In general, it is usually computed that a man weighing 70 kg. requires in calories:

2,500 for a sedentary life,
3,000 for light muscular work,
3,500 for medium muscular work,
4,000 and upwards for very hard toil.‡

‡McKillop.

These figures apply to the average man, but in calculating the calorie requirements of a family or a community we must make allowance for the lesser requirements of women and children. Several dietitians have compiled tables showing how many calories are expended according to age and sex, and the German authorities have recently taken these figures and from them calculated a generalized mean, which shows in comparison

with men the percentage that should be allowed for women and children. The figures are as follows:

Man	100
Woman	83
Boy over 16	92
Boy 14-16	81
Girl 14-16	74
Child 10-13	64
Child 6-9	49
Child 2-5	36
Child under 2	23

In calculating the calorie requirement of the population as a whole, the necessity of making allowance for the varying needs of men, women, and children would obviously make the calculations far too complicated for practical purposes. It is necessary to have a factor by which we may multiply the total population in order to determine its "man value." This factor is based on the relative proportion of men to women and children, and it amounts very nearly to 0.75, i. e., three-quarters of the total population gives "the man value." Knowing the total population, say, of a city, we must therefore multiply this by 0.75 in order to ascertain for how many men doing moderate muscular work (3000 C.) food has to be provided.

THE PROTEIN REQUIREMENT

The facts considered in the previous two chapters lead to the question: To what extent may the proportion of protein in the diet be reduced with safety? It is evident that there must be a minimum below which every one of the necessary building materials of protein could not be supplied in adequate amount to reconstruct the worn-out tissue protein.

The extent to which the protein content of the diet of man can be lowered with safety depends on several factors, of which the most important are: first, the nature of the protein; second, the number of non-protein calories; and third, the extent of tissue activity. Where so many factors must be taken into consideration, the only method by which the actual minimum can be determined consists in what may be called "cut and try experiments." Of the many investigations of such a nature, probably the best one for us to consider, is that recently published from the Nutrition Laboratory of Copenhagen. The subject, an intelligent laboratory servant, lived a perfectly normal and active life for a period of five months on a diet of potatoes cooked with margarine and a little onion, and containing 4000 C., with a total protein content of 29 grams. During another period he did outdoor work as a mason and laborer, and took 5000 C. daily, and 35 grams of protein.

It is important to contrast these results with the following based on municipal statistics of gross consumption.

MUNICIPAL FOOD STATISTICS

	PROTEIN	FAT	CARBOHYDRATES	CALORIES
	gm.	gm.	gm.	
Königsberg	84	31	414	2394
Munich	96	65	492	3014
Paris	98	64	465	2903
London	98	60	416	2665

It is certain that man can lead a normal existence and remain in good health on very much less protein than the 100 grams which statistical studies show to be the amount he actually takes. This discrepancy between the amount which experiment demonstrates to be adequate and that which habit and custom demand, raises the question as to whether, after all, our instincts may not have erred and so made us unnecessarily extravagant in our protein intake. It has been suggested that such protein extravagance will in various ways have a deleterious effect on the organism; thus, that the excretory organs, such as the kidneys, will be overtaxed in eliminating the unused amino acids, that the constant presence of these bodies in excess in the blood will cause degeneration and sluggish metabolism, and that the excess protein in the intestine will lead to the production of ptomaines, whose subsequent absorption into the blood will cause toxemic symptoms.

Important support to such views appeared to be supplied some dozen years ago by Chittenden, who was able to show that he himself and many other persons doing different kinds of work could be supported on daily amounts of protein that were not more than from one-third to one-half of the amount usually taken. Not only so, but it was averred that distinct improvement was experienced in the general sense of well-being and of mental efficiency as a result of the lesser protein consumption.

Taking these results as a whole, it is quite clear that man can get along under ordinary conditions with much less protein than he usually takes; but that really proves nothing, for the question is not *can* he but *should* he, so deprive himself? Are instinct and custom wrong and is Chittenden right? That is the question. To answer it many studies have been made of the condition of peoples who for economic or other reasons are compelled to live on less protein than the average. Are these people healthier, less prone to infections and degenerative diseases, and more efficient mentally than others? In such studies great care must be exercised to see that conditions other than diet, such as climate, exercise, etc., are properly allowed for. It would not be fair, for example, to compare the mental and bodily condition of peoples living in the tropics

and who take comparatively little protein, with those living in temperate zones, who consume much more. After discounting all of these other factors, it has been quite clearly shown that, when the protein allowance is *materially reduced,* the people as a whole are less robust, mentally inferior, and, instead of being less prone to the very diseases which are usually supposed to be due to overloading of the organism with useless excretory products, are more liable to suffer from them.

That a decided reduction in protein weakens the defense of the organism against infection is probably due to the fact that the fluids of the body normally contain a great variety of so-called antibodies—that is, of highly complex substances that are largely protein in nature. When bacteria, or the poisons produced by them, enter the body, they are met by one or more of these defense substances and destroyed or neutralized. Now it is clear that there should always be a surplus of protein-building materials from which the antibodies may be constructed. Such an excess will constitute a "factor of safety" against disease. And there are factors of safety of another nature to be provided for, two of which we are in a position to appreciate. In the first place, there must always be an adequate supply of tryptophane, of lysine, and of cystine, not only to meet the bare necessities of the protein constructive processes that go on under normal conditions, but also to make good the larger amount of protein wear and tear that greater degrees of tissue activity will entail. Although moderate muscular exercise does not appear to cause any immediate consumption of protein (carbohydrate and, later, fat being the fuel material used to produce it), yet it does throw a greater strain on the tissues and causes a greater wear and tear of the machinery, and hence a demand for more protein-building material. In the second place, there are certain of the internal secretions of the body, such as epinephrine (adrenaline), that are essential for life, and as crude materials for the manufacture of which certain amino acids are essential. Tyrosine is one of these, and since proteins, as we have seen, differ from one another quite considerably in the amount of this amino acid which they contain, it is advisable to provide an excess, so that an adequate supply of tyrosine may always be available.

The answer to one of the most important practical questions in dietetics—namely, "What proportion of protein should the diet contain?" depends on these scientific principles. The source of the protein is the important thing. With animal protein there is no doubt that we could get along with perfect safety by taking daily not more than 50 or 60 grams, which is about half of what we actually consume. If the protein is of vegetable origin and part of it of the first quality, as wheat and Indian corn preparations, more should be taken so as to allow for the

deficiency of certain amino acids. When vegetable proteins of the second quality, such as those of peas, beans, lentils, etc., are alone available, much larger amounts are necessary. Such proteins are inadequate in the case of growing children at least, and even in adults it is undoubtedly advisable that other proteins should supplement them.

To insure safety, therefore, it is almost imperative that the diet should contain *proteins of various sources.* If for economic reasons the main source must be proteins of vegetable origin, then some animal protein, such as is contained in milk or meat or eggs, should be added to at least one of the daily meals. When peas and beans are mainly depended on for the protein supply, they should be taken either with milk or one of its preparations, or with a thick gravy or sauce made from meat and containing the finely minced meat. This must not be strained off, for if it is, the sauce will contain only the meat extractives but not any of the protein, which is coagulated by the boiling water. Meat extract, in other words, contains no proteins; it is not a food but merely a condiment of no greater dietetic value than tea or coffee.

ACCESSORY FOOD FACTORS

Little need be added to what has already been said regarding this subject. The practical point to be remembered is that there are at least two accessory factors concerned, one of them soluble in fat and present in adequate amount in butter and other animal fats but not in vegetable oils, and the other soluble in water and present in wheat, vegetables, fruits, etc. Milk contains both of these factors, so that its inclusion in a diet is a safeguard not only against inadequacy in suitable protein, but also against the absence of accessory food factors. There is little danger of the diet being inadequate with regard to food factors if it contains some fruits or green vegetables or unheated fresh milk. The food factors are destroyed by prolonged cooking.

DIGESTIBILITY AND PALATABILITY

We have seen that practical dietetics depends on several factors, the exact relative importance of which can not perhaps be gauged in every case, but preparation of the food so as to make it appetizing must undoubtedly rank high. The importance of *good cooking* will now be apparent. It is the act of making food appetizing and therefore digestible. It is really the first stage in digestion, the stage that we can control, and one therefore to which much attention must be given, especially when it becomes necessary to make attractive articles of diet ordinarily considered common and cheap. Most people can cook a lamb chop so as to make it

reasonably appetizing, but few can take the cheaper cuts of meat and convert them into cooked dishes that are as popular and attractive. And there are still fewer who can take the left-overs and trimmings and convert them in the same way. This is the real art of cooking, and too much encouragement can not be given to the effort which our cooking experts are making to show people how these things can be done. The waste of good food in a large city is truly appalling.

Cooking has other advantages than making the food appetizing. The heat loosens the muscle fibers of the meat so that it is more readily masticated; it destroys microorganisms and parasites in the meat; it destroys antibodies which might interfere with the action of the digestive ferments. Thus, untreated raw white of egg is not digested in the stomach because it contains one of the antibodies which prevent the pepsin from acting on it; but boiled egg white, if properly chewed, can be digested, and whipping the egg white into a foam partly destroys the inhibiting substance.

Before concluding, something should be said about the *laxative qualities of food,* for it is often in this particular alone that one food is more satisfactory than another. A diet of meat, milk, eggs, and white bread is apt to be unphysiologic because there is nothing in it to act as what has been called intestinal ballast; that is, a material which will keep the intestines sufficiently filled to stimulate their muscular movements. This ballast is best furnished in the shape of cellulose, the most important constituent of green food. Peas, beans, cabbage, salad, and many fruits, especially apples, should always occupy a place in the daily menu. Another valuable food yielding this ballast is the outer grain of wheat, oats, etc. So much must not be taken as to produce a constant intestinal irritation, and each person must determine for himself where this limit lies. The difference among various breads is almost entirely in the degree to which they supply ballast.

The all-important subject of food economies can receive no attention here, except to point out that it is one which must be most carefully considered in the solution of all problems of dietetics. An admirable account of the subject will be found in Graham Lusk's "Science of Nutrition" (third edition) and in McKillop's "Food Values."[16]

CHAPTER LXVII

THE METABOLISM OF PROTEIN

Introductory.—The older physiologists believed that the protein taken with the food was brought into a soluble condition by the digestive enzymes, and that it was then absorbed into the blood and directly incorporated with the tissues. The discovery of the enzymes trypsin and erepsin and of free amino acids in the gastrointestinal contents clearly showed that this simple theory of Liebig could not be correct. It was, furthermore, found that when an excess of proteins such as egg albumin gains entry to the blood, part of the protein appears in an unchanged condition in the urine; and that enzymes capable of digesting this protein but not other varieties make their appearance in the blood.

After the injection of foreign proteins into the blood, symptoms of varying severity often develop, from the almost instantaneous death produced by snake venom to the slowly developing anaphylactic reactions which follow the injection into the blood of many proteins chemically indistinguishable from those of the blood serum itself. When protein is taken in the usual amounts by mouth, these poisonous reactions do not supervene,—even snake venom is harmless when swallowed,—nor is it possible during digestion of a protein meal to detect food protein in the blood by means of the precipitin reaction. Finally it was discovered that the very slow intravenous injection of completely digested flesh did not produce on the part of the body any of the reactions that injected protein itself produces, indicating that perfect assimilation had occurred. From these and similar observations it soon became clear that protein can not be absorbed as such from the alimentary canal, but must *first of all be completely broken down into the amino acids, which are then rebuilt into the protein of the organism.* The direct evidence for this important change in belief concerning protein metabolism has been gained by the discoveries that: (1) nitrogen equilibrium can be maintained in animals fed with completely digested protein mixtures; and (2) amino acids can be isolated from the blood.

The experiments of the first group consist, in principle, in breaking down protein until there is no longer the characteristic biuret test, and then feeding this digestion mixture to animals and observing them from day to day, using as criteria of their nutritional condition the body weight

and the nitrogen equilibrium. (Page 571.) It has been shown that success in maintaining nutritional efficiency depends partly on the nature of the process used for digesting the protein, and partly on the presence or absence of carbohydrate in the digestion mixture. It was found that acid hydrolytic products failed to maintain equilibrium, and it was believed that this was owing to the fact that the acid had more completely disrupted the protein molecule, and had left no polypeptides, which, it was imagined, remained intact during enzyme action and were essential for proper protein metabolism. This view has now been considerably altered, since it has been shown that the acid actually destroys certain amino acids which the enzyme leaves intact. The amino acid particularly concerned is tryptophane. Thus, when animals were fed with three diets, consisting of (1) fully digested casein, (2) fully digested casein from which the tryptophane had been removed, and (3) fully digested casein from which the tryptophane had been removed and then the proper amount of pure tryptophane added, it was found that nitrogen equilibrium could not be maintained on the second diet, which contained no tryptophane, whereas it was maintained on the first and third diets. That this explanation is correct is further supported by the fact that, if the protein is only partly digested by acid—that is, not digested enough so as to break up all the tryptophane—it can efficiently maintain nitrogen equilibrium.

Regarding the necessity for carbohydrates, it is possible that under certain conditions these may be produced from the protein itself. At least, it has been possible for Abderhalden, who has done a large share of this work, to maintain an animal in nitrogen equilibrium with a diet of digestion products and fat containing no carbohydrate.

These results obtained in different classes of animals have also been confirmed for the human subject. A boy suffering from a stricture of the esophagus, when fed by rectum for fifteen days with digestion products resulting from the action of trypsin and erepsin on flesh, gave evidence of nitrogen retention.

Concerning the second type of evidence, many investigators attempted to separate the amino acids themselves from the blood, particularly during the digestion of a large amount of protein, but the results were at first entirely negative because of the lack of methods that were sufficiently delicate to make it possible to detect the slight increase that could be expected even when a maximum absorption of nitrogen had occurred. The very large flow of blood through the portal vein causes such extensive dilution of any substances added to it that the concentration of the substance in an isolated sample of the blood can be only trivial.

To account for the indisputable disappearance of the amino aci the intestine during protein digestion, coupled with the impossib detecting any of them in the blood, two views were current fo years. One of these was that the amino acids become deaminate group split up as NH_3) by the intestinal epithelium, and the oth these cells are endowed with the power of reconstructing the amir into protein, which then passes into the blood. Justification for amination hypothesis seemed to be obtained by the observation th is more free ammonia in the blood of the portal vein than in tha systemic circulation. The falsity of this evidence was, howeve nitely established by Folin and Denis,[32] who found by means of quantitative methods for the estimation of ammonia and urea in th that the amount of neither of these substances became increased portal blood during absorption of amino acids from the intestine made the further important discovery that the ammonia in the blood is really very little in amount, and represents that abso such from the intestinal lumen, where it is produced chiefly by th of putrefactive bacteria.

Nor could any evidence be obtained in favor of the hypothe the absorbed amino acids become built up in the intestinal epi into proteins, which are then transformed or carried away by th This hypothesis was based entirely on negative findings, and ha fore to be dropped when discovery was made of the actual pres amino acid in the blood.

This brief historical survey of the subject brings us to a positio we may proceed to discuss the present-day teaching regarding metabolism. Briefly stated, this teaching is to the effect that *the molecule is broken down into its ultimate building stones, the amin by the digestive enzymes of the gastrointestinal tract, and that thes acids are absorbed into the blood, by which they are carried to the organs and tissues, which sift out the amino acids and use those which they require for the reconstruction of their broken-down The amino acids not required for the process, along with those wh be liberated in the tissues themselves by disintegration of tissue are then split into two portions, one represented by ammonia and tl by the remainder of the amino acid molecule. The former is exc urea and the latter is oxidized to produce energy.*

CHEMISTRY OF PROTEIN

Before proceeding to discuss the evidence upon which the abo clusions depend, it will be necessary to consider some of the most im facts concerning the chemistry of the protein molecule. We shall

this information not only to understand the history of protein in the animal body, but also to follow intelligently the important work that has already been discussed concerning the relative value of different proteins as food. A knowledge of protein chemistry has come to be essential in practically all branches of medical science.

Proteins, like starches, are composed of numerous smaller molecules. In the case of starch these molecules are the various monosaccharides—glucose (dextrose), levulose and galactose; in the case of proteins they are the amino acids. The breaking apart of the links that hold the molecules together is effected in both cases by the process of hydrolysis, so called because of the fact that the reaction consists in the taking up of a molecule of water at each of the places where the chain falls apart. This hydrolysis may be effected either by the action of mineral acids or alkalies, or by enzymes, the only difference in the action of these reagents being that in the former case the breaking apart takes place more or less indiscriminately, whereas in the latter it proceeds according to a definite plan, which varies somewhat with the type of enzyme employed. Just as a chemical knowledge of the structure of sugar or monosaccharides is the basis of carbohydrate chemistry, so is that of the amino acids the basis of protein chemistry.

Amino Acids.—There are, so far as known, eighteen different amino acids concerned in the constitution of protein, but they are all alike in their characteristic structure. The most striking characteristic depends on the presence in the molecule of: (1) an amino group with a basicity comparable to that of ammonia, and (2) an acid group with an acidity comparable to that of acetic acid. Let us take in illustration one of the simplest fatty acids—namely, acetic. It has the formula CH_3COOH. The COOH group is called *carboxyl*, and on it depend the acid properties of the compound. The CH_3 group is known as *methyl*, and the *amino group* (NH_2) is attached to it in place of one of the hydrogen atoms, thus giving the formula CH_2NH_2COOH, which is aminoacetic acid or glycocoll. If we take the next higher acid of the fatty acid series, having the name propionic and the formula CH_3CH_2COOH, its amino acid, called alanine, has the formula CH_3CHNH_2COOH. Now let us place the formulas of these two acids side by side in the following manner:

$$\begin{array}{ccc} & H & \\ & | & \\ NH_2 - & C & - COOH \\ & | & \\ \text{(amino group)} & H & \text{(acid group)} \end{array} \qquad \begin{array}{ccc} & CH_3 & \\ & | & \\ NH_2 - & C & - COOH \\ & | & \\ \text{(amino group)} & H & \text{(acid group)} \end{array}$$

Aminoacetic acid (glycocoll) — Aminopropionic acid (alanine)

It will be observed that the only difference between the two acids is dependent upon a change in the group that is attached to the upper vertical valency bond of the central carbon atom, which therefore must be considered as the center of the entire molecule. The various amino acids produced from protein differ from one another solely with regard to the chemical nature of the group that is attached to this vertical valency bond. Evidently, then, *the reactions that amino acids possess in common* must depend on the terminal groups containing the carboxyl and amino radicles, whereas *the characteristic reaction of each of the eighteen amino acids* must depend upon the differences in the radicles attached to the upper vertical bond. This may be represented thus:

$$\begin{array}{c} \text{Any radicle} \\ | \\ NH_2-C-COOH \\ | \\ H \\ \text{Any amino acid} \end{array}$$

The end groups endow the amino acids with the power to combine with both acids and bases. With acids they behave like substituted ammonias to form salts, which can ionize into the amino acid, as the cation, and the acid group, as the anion. With bases the carboxyl group reacts to form salts, which yield amino acid as the anion. A most important reaction consists in *the condensation of aldehydes with the amino group*. This occurs particularly readily with formaldehyde, water being eliminated in the reaction, and the basic nature of the amino acid being thus destroyed. Upon this reaction depends the method of Sörensen for determining the amount of amino acid in a mixture (see page 606). The titration is performed by rendering the solution of amino acids neutral, then adding formaldehyde and titrating with standardized acid, using phenolphthalein as the indicator, and thus finding to what degree the acidity of the mixture has become increased as a result of adding the formaldehyde. Since this increase in acidity must depend upon the number of amino groups, it furnishes us with an indirect estimate of the concentration of the amino acids. The reaction is illustrated by the equation:

$$\underset{\text{(amino acid)}}{\begin{array}{c} \text{radicle} \\ | \\ NH_2-C-COOH \\ | \\ H \end{array}} + \underset{\text{(formaldehyde)}}{\begin{array}{c} H \\ | \\ H-C=O \\ \\ \end{array}} = \begin{array}{c} \text{radicle} \\ | \\ CH_2=N-C-COOH \\ | \\ H \end{array} + H_2O$$

Another reaction of amino acid of physiologic interest is that known as the *carbamino reaction*, consisting in a union of the amino acid with calcium and carbonic acid. Finally, it is important to note that the amino

group is very firmly attached; it remains intact in acids and alkalies and is removable only by a process of oxidation. This can be accomplished by treating the amino acid with such reagents as hydrogen peroxide or permanganate, when the amino group is displaced and a so-called ketonic acid formed. The reaction will be evident from the accompanying equation:

$$\begin{array}{ccc} \begin{array}{c} CH_3 \\ | \\ O + NH_2-C-COOH \\ | \\ H \end{array} & \rightleftarrows & \begin{array}{c} CH_3 \\ | \\ O=C-COOH + NH_3 \\ \\ \\ \end{array} \\ \text{(alanine)} & & \text{(pyruvic acid)} \end{array}$$

To illustrate this reaction we have chosen aminopropionic acid or alanine, because the substance formed by its oxidation and known as pyruvic acid is of very great importance in intermediary metabolism. It serves as the common substance from which proteins, carbohydrates or fats may be formed, and therefore as the intermediary substance through which one of them may pass on being transformed into another (page 666). The use of two arrows pointing in opposite directions in the above equation indicates that the reaction may proceed readily in either direction.

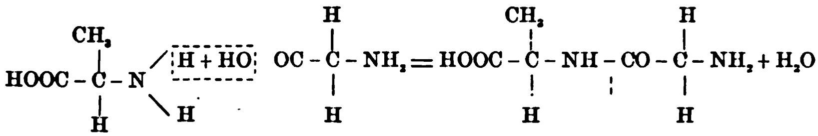
The ammonia set free from amino acids may be oxidized to free nitrogen by using nitrous acid according to the general equation: $NH_3+HONO=2H_2O+N_2$. Upon this reaction depends another extremely important *quantitative method for measuring the number of amino groups present in protein* (*Van Slyke*). To make the estimation, nitrous acid is allowed to act on the amino acids, and the volume of nitrogen gas set free by the reaction is measured, the principle being similar to that used for the determination of urea by the hypobromite method.

The apparatus employed for decomposing the substance and collecting and measuring the evolved nitrogen consists essentially of a mixing bulb, connected below through stopcocks with two small burettes, one containing a solution of sodium nitrite and glacial acetic acid, and the other a solution of the substance to be investigated. The upper end of the mixing bulb is connected through a three-way cock with a graduated gas burette and with another bulb containing potassium permanganate solution. By allowing some nitrite and acid solution to run into it and shaking, the mixing bulb is first of all filled to a certain mark with nitrous oxide gas. A measured quantity of the amino solution is then allowed to mix with the nitrite; the apparatus is shaken for five minutes at 15 to 20° C., and the evolved nitrogen and nitric oxide are driven over into the permanganate, which absorbs the nitric oxide, leaving the nitrogen, which is then measured in the burette.

The apparatus has now been so perfected that numerous analyses may be made with it in a very short time and with a degree of accuracy that is scarcely surpassed in any other biochemical estimation.

From the point of view of protein chemistry, the most significant reaction of the amino acids is their ability to link together to form compounds

called *peptides*. This linking occurs between the amino group of one amino acid and the carboxyl group of the other. When alanine and glycocoll, with which we are familiar, are thus linked together, the reaction takes place according to the equation:



(alanine) (glycocoll) (alanyl - glycocoll)

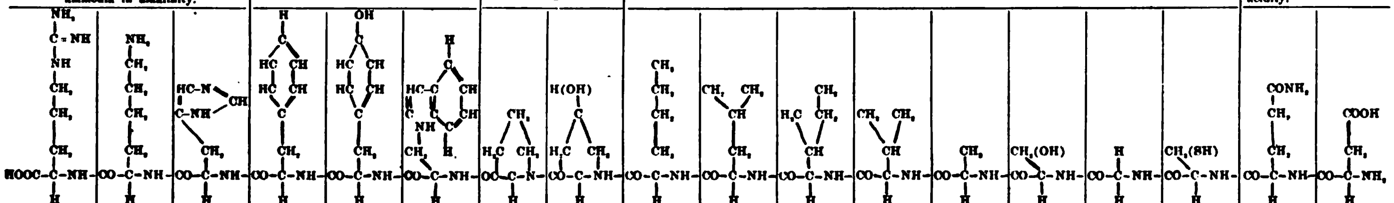
In this manner, then, a so-called dipeptide is formed, in which it will be observed there still remain free carboxyl and amino groups, thus permitting the linking on of other amino-acid groups to form tripeptides or tetrapeptides or other polypeptides. Indeed, this process of condensation may go on practically indefinitely, a polypeptide containing eighteen amino-acid groups—namely, three leucine and fifteen glycocoll groups—having actually been synthesized. The resulting polypeptides have the properties of derived proteins like the proteoses; thus, they give the biuret and other reactions characteristic of proteins and are precipitated by such reagents as mercuric chloride and phosphotungstic acid. Some are also digested by trypsin and erepsin. They have the same optical activities as proteins. One of them has been prepared which produces a typical anaphylactic reaction. So far a polypeptide that can be coagulated by heat or is in other ways identical with the naturally occurring proteins, has not been synthesized; but there is no doubt that it is only a matter of time before this will be accomplished.

Eighteen distinctly different amino acids have been isolated from protein, and it may assist in the conception of our problem if we place these amino acids side by side and link them together in the manner described above. This is done in the accompanying chart compiled by D. D. Van Slyke, in which also various other important facts concerning the chemistry of the amino acids are incidentally added.

At the lower part of each formula will be seen the characteristic carboxyl and amino groups of neighboring acids linking together the terminal carbon atoms. The upper vertical bond of this carbon atom is connected with the characteristic group of the amino acid, which may be very simple and represented only by hydrogen, as in glycocoll, or highly complex and including a ring formation, as in tryptophane. It will further be observed that *there may be other amino groups* connected in various positions in this radicle. This is particularly the case in the first three of the amino acids in the table—namely, the basic amino acids. In lysine the extra amino group reacts with nitrous acid, liberating free nitrogen

Formula of an Imaginary Protein Containing One Molecule of Each Known Amino Acid

	Basic Amino Acids Two or more basic groups to one acid COOH group. The three "hexone bases." Resemble ammonia in alkalinity.			Neutral amino acids contain one amino group and one carboxyl group, which neutralize each other													*Acid Amino Acids* Two acid COOH groups to one NH_2 group. Resemble acetic acid in acidity.	
				Aromatic, containing benzene rings			Containing Pyrrolidine rings		Containing only aliphatic chains straight or branched									
	Arginine	Lysine	Histidine	Phenyl-alanine	Tyrosine	Trypto-phane	Proline	Oxypro-line	Nor-leucine	Leucine	Iso-leucine	Valine	Alanine	Serine	Glycocoll	Cysteine	Glutamic acid amide	Aspartic acid
Form glucose	+	-	-	-	-	-	+	?	?	-	-	-	+	+	+	+	+	+
Form acetone	-	-	+	+	+	-	-	?	?	+	?	-	-	?	-	?	-	-



(Van Slyke: Harvey Lectures, 1915-1916, Page 148.)

by the Van Slyke method; but in other cases, as in arginine, it fails to give this and the other characteristic reactions of the amino group. That the extra amino group in lysine reacts directly with nitrous acid explains why various proteins when examined for amino nitrogen yield an amount that is equal to half of the lysine nitrogen.

It will further be observed that the amino acids are arranged in three main groups: one basic, another neutral, and the third acid. *The acids of the basic group* are three in number and have an alkalinity similar to that of ammonia. They have been called the *hexone bases,* because each contains six carbon atoms. They are alone present in certain forms of protein called protamines. *The neutral amino acids* contain one amino group and one carboxyl group, which exactly neutralize each other. This is the largest group of amino acids, and is further subdivided into three: one containing aromatic or benzene rings and including the very important amino acids, tyrosine and tryptophane; another containing the so-called pyrrolidine ring; and the third, the largest of all, containing the so-called aliphatic chains; that is, the chains characteristic of the fatty acids and which may be either straight or branched. When the chains are branched, the substance is called an isosubstance, as in isoleucine. *The acid amino acids*, including glutamic acid and aspartic acid, are characterized by containing two carboxyl groups and only one amino group. They therefore resemble acetic acid in acidity.

It may be of assistance to some if we restate these chemical facts from a slightly different standpoint as follows:

Glycine, or glycocoll, is aminoacetic acid, CH_2NH_2COOH.

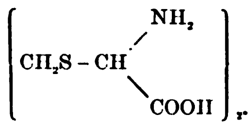
Alanine is glycine *plus* a methyl group, $CH_3CH{<}^{NH_2}_{COOH}$; it is therefore aminopropionic acid and is closely related to lactic acid, which is $CH_3CH{<}^{OH}_{COOH}$. *Many of the other amino acids may be considered as derivatives of alanine,* thus:

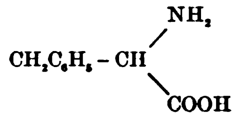
1. *Serine* is alanine with an "OH" (hydroxyl) group in place of one of the "H" atoms of the methyl group, $CH_2OH-CH{<}^{NH_2}_{COOH}$.

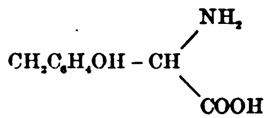
2. *Cysteine* is alanine with an "SH" (thio) group in this position,

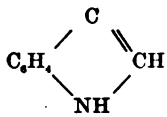
$$CH_2SH-CH{<}^{NH_2}_{COOH} .$$

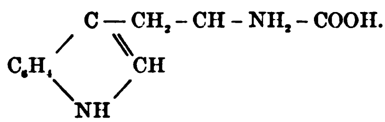
Two cysteine molecules united at the "S" groups give *cystine*.

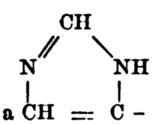


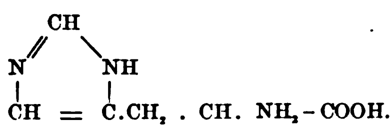
3. *Phenylalanine* has a C_6H_5 (phenyl) group, .

4. *Tyrosine* has a C_6H_4OH (phenol) group, 

5. *Tryptophane* has a  (indole) group:

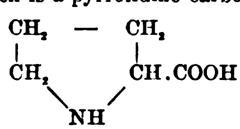


6. *Histidine* has a  (imidazole) group:



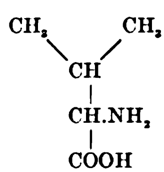
The last two are also called *heterocyclic compounds,* of which there is another, viz.;

Proline (and oxyproline), which is α-pyrrolidine carboxylic acid:

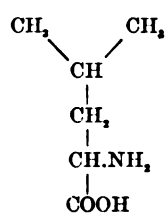


Other amino acids are:

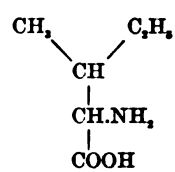
(1) Valine
Leucine
Isoleucine
thus:



(valine)



(leucine)



(isoleucine)

(2) The amino dibasic acids:

Aspartic, which is aminosuccinic acid,

$$\begin{array}{l} CH_2COOH \\ | \\ CHNH_2COOH \end{array}$$; and

Glutaminic, which is aminoglutaric acid,

$$\begin{array}{l} CH_2 \\ | \\ CH_2 - COOH \\ | \\ CHNH_2\ COOH. \end{array}$$

Lastly there are the diamino acids, in which two groups exist:

Lysine a ε-diaminocaproic acid,

$$NH_2CH_2 - CH_2 - CH_2 - CH_2 - CH_2 - CH \begin{array}{l} \diagup NH_2 \\ \diagdown COOH. \end{array}$$

Arginine α-amino—δ-guanidine-valerianic acid,

$$HN = C \begin{array}{l} \diagup NH_2 \\ \diagdown NH.CH_2 - CH_2 - CH_2 - CH \begin{array}{l} \diagup NH_2 \\ \diagdown COOH \end{array} \end{array} .$$

The guanidine group in this acid is of interest because of its close relationsl

urea, which is $O = C \begin{array}{l} \diagup NH_2 \\ \diagdown NH_2 \end{array}$.

CHAPTER LXVIII

THE METABOLISM OF PROTEIN (Cont'd)

AMINO ACIDS IN THE BLOOD AND TISSUES

In the Blood.—Furnished with the general facts concerning the chemistry of proteins, we may now proceed to consider the more precise knowledge recently acquired concerning the history of this substance in the animal economy. Although no one has succeeded in separating amino acids in pure condition from drawn blood even during the height of digestion, it has nevertheless been possible to do so from circulating blood by a method of dialysis, known as *vividiffusion,* elaborated by Abel[33] and his pupils. The method consists in connecting a long tube of collodion with the two ends of a cut artery in an anesthetized animal. The tube, coiled many times, is then immersed in a solution containing approximately the same salt content as the blood plasma of the animal. The diffusible constituents of the blood plasma dialyze into the saline solution; or any one of them may be prevented from dialyzing by adding that particular substance to the saline in such amounts as will make its concentration in plasma and saline alike. In some ways, it will be seen, the apparatus may be considered as an artificial kidney. Its possible clinical application for the purpose of removing poisons from the blood is under investigation. It has been possible in this way to isolate several of the amino acids and other ammonia-yielding substances from blood. Thus, alanine and valine have been obtained as crystalline salts, and histidine and creatine (see page 622) shown to be present by their reactions. All of the amino substances, however, do not dialyze, and these exceptions are further characterized by the fact that they do not readily give up their ammonia on the addition of sodium carbonate, as do the diffusible substances (Rohde). Although amino acids can thus be separated in a pure state from circulating blood, their concentration in a drawn specimen is too low to make direct quantitative estimation possible. By the methods of Van Slyke and Sörensen, already described, however, it has been shown among other things that the blood always contains a certain concentration of amino acids; thus, in that of fasting animals from 3 to 5 mg. per 100 c.c. of blood are usually found present. During the absorption of a protein meal, the amino content of the blood undergoes

a marked increase, becoming doubled or more; and a similar result has been obtained by placing pure amino acids in the small intestine. After 10 grams of alanine, for example, the amino nitrogen of the mesenteric blood rose from 3.7 to 6.3 mg. per cent.*

In the Tissues.—After entering the circulation, the amino acid *very quickly disappear* from it again. This has been demonstrated by observing the amount of amino acids in the blood after intravenously injecting a solution of amino acid into an anesthetized animal. After injecting 12 gm. of alanine into the vein of a dog, 90 per cent was found

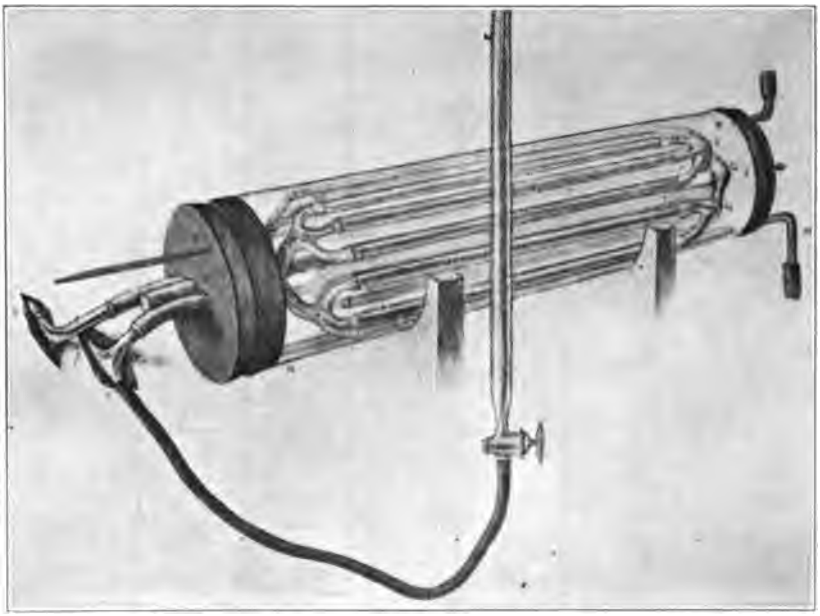


Fig. 186.—Vividiffusion apparatus of J. J. Abel.

to have disappeared from the circulation within five minutes. The question is, What becomes of the amino acids that rapidly disappear? Are they decomposed in the blood, or do they become absorbed by the tissues? This problem has been attacked by analyzing portions of various organs and tissues removed before and some time after the injection into an animal of amino acid solutions. In the case of the muscles it has been found that the amino-acid content increases until from 60 to 80 mg. per cent of amino acid has accumulated. Beyond this point, however, the muscles do not seem to be able to take up any more amino acid. The capacity of the intestinal organs, however, is more elastic;

*This is a convenient way of stating per 100 c.c. of blood.

for example, the amino nitrogen of the liver has been observed to become increased to 125 or 150 mg. per cent of the original amount. Although this absorption of amino acids by the tissues is extremely rapid, it never proceeds to such a point that the blood becomes entirely free of them. Even after many days' starvation the blood contains its normal quota of from 3 to 10 mg. per 100 gm. of moist tissue (Fig. 188). This indicates that a certain equilibrium must become established between the amino-acid content of the blood and that of the tissues, the concentration in the tissues being approximately from five to ten times greater than in the blood.

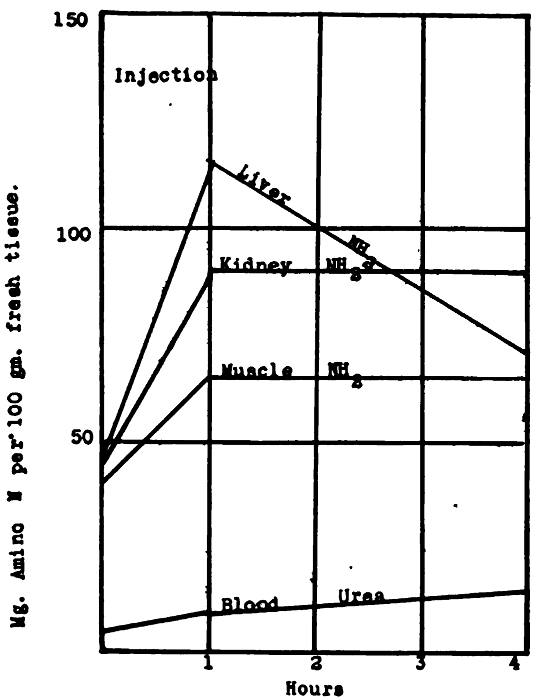


Fig. 187.—Curves showing the amount of amino nitrogen taken up by different tissues after the cutaneous injection of amino acids. The lowermost curve shows the urea concentration of the blood. (From D. D. Van Slyke.)

The absorbed amino acids are very loosely combined with the tissues, for they can be extracted by such feeble reagents as water or dilute alcohol. Their presence can not, however, be merely due to diffusion; for if it were, the concentration could not become greater in the tissues than in the blood. The further fate of the amino acids is difficult to follow. We know that they do not remain in the body for a long time, because most of the protein nitrogen in the food is excreted as urea within twenty-four hours after ingestion; and when single amino acids are fed, they quickly reappear in the urine as urea.

The tissues can therefore be only a stopping-place for the acids. When the latter are determined in blood collected from di parts while absorption of protein from the intestine is in proc has been found, as shown in Fig. 188, that during the passage blood through the liver there is a greater fall in the concentrati amino acids than during its passage through the entire remainc the body.

It will be seen that the above conclusions are drawn from e tions made on blood taken from the vena cava, portal vein, and h

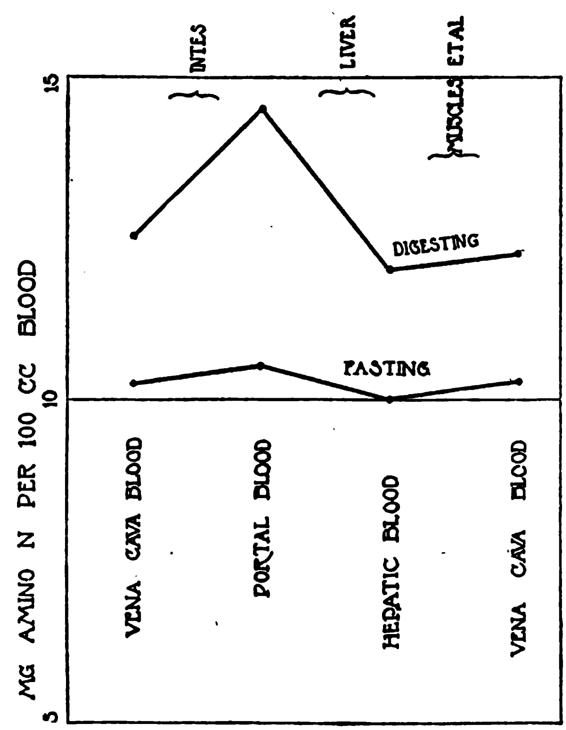


Fig. 188.—Curves showing the concentration of amino-acid nitrogen in the blood durin and protein digestion. (From D. D. Van Slyke.)

artery, the upper curves in the chart being from animals during dig and the lower from fasting animals. The results show that the live be particularly greedy of amino acids, which, however, must rapid come transformed into other substances, since no conspicuous tion has been found to occur in the amino-acid content of the according to whether the animal is fasting or is digesting protein This result, it is to be noted, is quite different from that which tained after the intravenous injection of amino acids, and the res

the two experiments taken together, indicate that the amino acids after their absorption can not remain in the tissues in a free condition for a long time. It means that *the amino acids during natural digestion must be disposed of at a rate which is practically the same as that at which absorption is proceeding.*

THE FATE OF THE AMINO ACIDS

To follow the metabolism of the amino acids further we must determine the end product into which they are converted. This is **urea**, whose estimation can nowadays be made with considerable accuracy on account of the discovery, by Marshall, of the action of urease in converting its nitrogen into ammonia, which can then be estimated by comparatively simple methods (Folin).

When the viscera are compared before and at various periods after the intravenous injection of amino acids, the immediate increase in amino nitrogen remains undiminished in all of them except the liver, in which a very rapid reduction is observed to occur. At the same time the percentage of urea in the blood steadily rises. These facts are illustrated in Fig. 187.

The simplest interpretation of these results is that the liver converts the amino acids into urea and discharges this urea into the blood. This conclusion, however, it must be observed, is not inevitable; for it is possible that the amino acids may be condensed into polypeptides in the liver, just as sugar is condensed by this organ into glycogen, and that the increase in urea is merely coincident (Fiske).

It must not be imagined that the conversion of the amino acids into urea is exclusively a function of the liver. On the contrary, it is well known that this process may occur in animals from which the liver has been entirely removed. It is probably safe to conclude, however, that the liver is the most active center for amino-acid transformation and urea formation.

When urea is estimated in samples of blood removed at short intervals of time after the ingestion of a large amount of protein, it is found that the increase becomes very early established. In one experiment, before the food was taken the concentration of urea nitrogen in the blood was a little over 10 mg. per cent; one hour after taking 500 grams of meat, it had risen to about 18, and in two hours to nearly 25. Evidently the increase had occurred about the same time as the passage of food from the stomach into the duodenum. These facts indicate that urea formation in the liver becomes stimulated long before the other tissues, such as the muscles, have had time to take up their full quota of amino

acids. During digestion of protein the liver does not appear to wait until the other tissues have become saturated with amino acids before it begins to destroy the unnecessary excess by conversion into urea; on the contrary, this process sets in with the very first installment of amino acid that reaches the liver by the portal blood. This conclusion is in harmony with the well-established fact that, when protein is given to a starving animal, the greater part of its nitrogen is soon excreted as urea, leaving only a small fraction to be used for rebuilding the wasted tissues (see page 643).

The amino acids that are absorbed by the extrahepatic tissues become very quickly converted into formed protein, as is evident from the fact that the concentration of free amino acids in the tissues of an animal during absorption of protein is not perceptibly greater than in those of a fasting animal, and the question remains to be considered, *What becomes of the protein thus formed?* The answer is, that it is gradually used up in the metabolic processes, so as to liberate again the amino acids, which add themselves to those absorbed from the intestine and become used again or excreted, according to the demands of the tissues at the time for amino acid.

This process of liberation of amino acid from the breakdown of body protein goes on of course irrespective of absorption of amino acid from the intestine. It goes on, for example, during starvation; indeed, in this condition the percentage of free amino acids in the muscles is, if anything, somewhat higher than that observed in an ordinarily fed animal. In starvation also the migration of amino acid is going on among the various organs, of which those whose activity is essential to the maintenance of life, such as the heart and the respiratory muscles, are supplied with amino acids from tissues that are less vital, such as the skeletal muscles (see page 568). These experiments further show that free amino acids can not serve to any significant extent as food reserves in the same way as glycogen and fat. If amino acids were of value as food reserves, we should expect the store of them to be depleted by starvation. As to how long a period of time elapses between the incorporation of the absorbed amino acids into tissue protein and their subsequent liberation again by autolysis, we are entirely ignorant.

The researches which we have just been considering do not throw any light on the *relative value of different proteins in tissue metabolism.* They do not inform us as to which of the amino acids must be absorbed ready-made from the digested food, and which of them may be dispensed with since the organism can manufacture them for itself. We know that the higher animals can synthesize some amino acids, such as glycocoll, but not others, such as tryptophane; but which amino acids belong to

the glycocoll and which to the tryptophane groups, can not as yet be definitely stated. The investigation of this problem has to be undertaken by experiments of an entirely different type—namely, by observing the welfare and growth of animals fed on proteins of varying amino-acid composition. A full discussion of these experiments is given in the chapters on Nutrition and Growth.

CHAPTER LXIX

THE METABOLISM OF PROTEIN (Cont'd)

THE END PRODUCTS OF PROTEIN METABOLISM

Introductory.—So far we have approached the problem of protein metabolism by studying the behavior of the absorbed products of protein breakdown, and we have seen that these become gradually assimilated by the tissues and used by them in their metabolic processes. We have been unable, however, to offer any facts regarding the exact chemical changes which each amino acid undergoes during this process of tissue metabolism. At first sight it might appear an easy matter to collect such information by direct examination of the tissues themselves, either by searching in them for amino derivatives which might be derived from absorbed amino acids, or by studying the changes which occur when the amino acids are subjected to the action of the isolated tissue enzymes that must be responsible for the change. Such methods of investigation are, however, fraught with technical difficulties so great that very little can be learned from them, and for the present at least we must be content to piece our information together from facts derived by less direct methods. Such a method is offered by investigating the behavior of the end products of protein metabolism.

The main end product is *urea* along with traces of its precursor *ammonia,* but these are not the only ones, for some amino acids after being incorporated with the tissue proteins break down into products that are no longer members of the amino-acid series, although they may be closely related to certain amino ocids. Such substances are *creatine* and its anhydrid *creatinine.* A part of the amino acids during their presence in a free state in the blood may also be excreted unchanged by the kidney. Our list so far therefore includes urea, ammonia, creatine, creatinine, and amino nitrogen, of which the last is usually included in metabolism investigations in the fraction designated *undetermined nitrogen.*

Another group of closely related substances coming, not from the general protein metabolism of the tissues, but from the metabolism which is peculiar to the nuclei, consists of the so-called *purine bodies.* Furthermore, so as to serve as a check on results obtained by examining these nitrogenous metabolites, it is important to observe the manner of

excretion of the sulphur moiety of the protein molecule, for it will be remembered that it is in protein alone that sulphur is usually taken into the animal body. The excretion of *sulphur* therefore runs more or less parallel with the intensity of protein metabolism.

After selecting the end products that are most likely to be of significance, the first question concerns the amount of each of them excreted during twenty-four hours on diets that are either rich or poor in protein. The possibility of conducting such investigations obviously depends on the use of quick and yet reliable methods for the estimation of the nitrogenous metabolites. Such methods have been furnished by the painstaking and careful work of Folin, an example of whose results are given in the accompanying table.

	NITROGEN-RICH DIET	NITROGEN-POOR DIET
Volume of urine	1170 c.c.	385 c.c.
Total nitrogen	16.8 grams	3.60 grams
Urea nitrogen	14.7 grams = 87.5%	2.20 grams = 61.7%
Ammonia nitrogen	0.49 gram = 3.0%	0.42 gram = 11.3%
Uric-acid nitrogen	0.18 gram = 1.1%	0.09 gram = 2.5%
Creatinine nitrogen	0.58 gram = 3.6%	0.60 gram = 17.2%
Undetermined nitrogen	0.85 gram = 4.9%	0.27 gram = 7.3%
Total SO_3	3.64 grams	0.76 gram
Inorganic SO_3	3.27 grams = 90.0%	0.46 gram = 60.5%
Ethereal SO_3	0.19 gram = 5.2%	0.10 gram = 13.2%
Neutral SO_3	0.18 gram = 4.8%	0.20 gram = 26.3%

(Folin.)

The general conclusions which may be drawn from these results are as follows:

1. With a protein-rich diet much more urine is excreted in twenty-four hours than with one that is protein-poor. Evidently the nitrogenous metabolites act as diuretics.

2. The total or absolute amounts of nitrogen and of all the other nitrogenous metabolites, save creatinine, become diminished during the starvation period. The same is true of the sulphur derivatives, except in the case of the neutral sulphur, which behaves like creatinine.

3. The decrease in the portion of nitrogen excreted as urea is relatively greater than the decrease in total nitrogen, this fact being shown in the table by the percentage figures, which were secured by calculating the proportion of nitrogen in the various substances as a percentage of the total nitrogen excreted during the periods. The inorganic sulphate behaves in a manner similar to the urea—that is, the percentage of total sulphate excreted in the inorganic form becomes much less during starvation.

4. The relative amount of all the other nitrogenous metabolites, as well as that of the ethereal and neutral sulphates, becomes increased during starvation.

The most striking results of the above investigation are that creatinine remains unchanged during starvation, but that urea becomes relatively increased. The former must be derived from metabolic processes going on in the tissues independently of the supply of foodstuff carried to them, whereas the latter must depend, if not entirely, yet very largely, on the protein content of the food. Creatinine may therefore be called an end product of *endogenous* metabolism, and urea an end product of *exogenous* metabolism.

Other metabolites—namely, ammonia, uric acid and the undetermined nitrogen, as well as the ethereal sulphates—must represent processes of metabolism that are partly exogenous and partly endogenous.

Having made ourselves acquainted with the general nature of the changes that occur in the nitrogenous metabolites when protein metabolism is stimulated by the taking of food or depressed by starvation, we may now proceed to take up each of the metabolites separately and see what other information can be obtained regarding their source and origin in the animal body.

UREA AND AMMONIA

For various reasons it is important to consider these two metabolites together. During the intermediary metabolism of the majority of the amino acids, the amino group becomes broken off as ammonia, which immediately combines with the available acids to form neutral ammonium salts. The most available acid for this purpose is carbonic acid; therefore ammonium carbonate is formed in large amounts. A small proportion of the ammonia may combine with other acid radicles, such as chlorine, to form ammonium chloride. The fate of these two types of salt is very different. The ammonium carbonate becomes quickly transformed into urea, whereas the ammonium chloride is excreted in the urine. The process of urea formation may therefore be considered as having the function of preventing the accumulation of ammonium carbonate in the animal body. It is the means by which a harmful substance is converted into an innocuous substance—a detoxication process, in other words.

Regarding *the nature of the chemical process* involved in this transformation of ammonium carbonate into urea, reference to the formulas below will show that the ammonium carbonate that is formed by the union of carbonic acid with ammonia, by losing one molecule of water becomes ammonium carbamate, which by repetition of the process becomes transformed into urea:

$$\underset{\text{(carbonic acid)}}{CO\begin{matrix}\diagup OH\\ \diagdown OH\end{matrix}} \underset{\text{(ammonia)}}{+2NH_3} \rightleftarrows \underset{\text{(ammonium carbonate)}}{CO\begin{matrix}\diagup ONH_4\\ \diagdown ONH_4\end{matrix}} - H_2O \rightleftarrows \underset{\text{(ammonium carbamate)}}{CO\begin{matrix}\diagup ONH_4\\ \diagdown NH_2\end{matrix}} - H_2O = \underset{\text{(urea)}}{CO\begin{matrix}\diagup NH_2\\ \diagdown NH_2\end{matrix}}$$

Some of the urea may come from metabolic processes of an entirely different type. One of these at least is known; namely, the splitting-off of urea from arginine, which it will be remembered is guanidine-amino-valerianic acid (see page 605). An enzyme called arginase, having this action, has been isolated from various organs and tissues. The diamino-valerianic acid, or ornithine, which remains after the urea is split off, may be further used in protein metabolism. The reaction is shown in the following equation:

$$\underset{\text{(arginine)}}{NH_2-\underset{\underset{NH}{\|}}{C}-NH-CH_2-CH_2-CH_2-CHNH_2-COOH} + H_2O$$

$$= \underset{\text{(urea)}}{\begin{matrix}NH_2-CO\\ |\\ NH_2\end{matrix}} + \underset{\text{(ornithine)}}{NH_2-CH_2-CH_2-CH_2-CHNH_2-COOH}$$

On an ordinary diet, as we have seen, a man excretes somewhat more than 90 per cent of his total nitrogen as urea and about 3 per cent as ammonia, the remainder of the nitrogen appearing in the other nitrogenous metabolites.

Influence of Acidosis on Ammonia-Urea Ratio.—It sometimes happens that a large proportion of the ammonia is not converted into urea, but is used for the purpose of neutralizing abnormal acids present in the organism. When mineral acids are given to an animal, or when acids are produced in the organism itself by some faulty type of metabolism, the ammonia excretion by the urine immediately rises. In diabetes, for example, where considerable quantities of β-oxybutyric acid are produced (see page 683), a decided increase in the ammonia excretion by the urine is observed. A milder type of acidosis may also be induced in normal persons by withholding carbohydrates from the diet, and here again the ammonia excretion is relatively increased.

In such cases it is quite evident that ammonia is used as *an alkaline reserve* of the body; that is, as a substance which is capable of preventing acidosis by neutralizing the acids. It does not appear, however, that all types of acidosis entail the utilization of ammonia as reserve alkali, and an increase in the relative amount of ammonia in the urine does not necessarily indicate a condition of acidosis. In the pernicious

vomiting of pregnancy, for example, a relatively high excretion of ammonia has been found associated with no greater a degree of acidosis, as determined by the power of the plasma to absorb carbonic acid, than in normal cases of pregnancy.

Influence of Liver on Ammonia-Urea Ratio.—Experimental Observations: (1) REMOVAL OF LIVER.—There are several facts which indicate that other causes than acid-production may interfere with the conversion of ammonia into urea. What are these causes? Since, as we have seen, the liver is the organ which most actively converts amino acids into urea, it would be natural to expect that, when the functions of this organ were interfered with, relatively more of the nitrogen excretion would occur as ammonia and relatively less as urea. In order to determine the exact significance of the liver as a urea-forming organ, two types of investigation have been used; namely, (1) observation of the changes produced in the ammonia-urea ratio in the urine by partial or total removal of the liver; and (2) observation of the urea-forming power of a liver perfused outside the body.

To remove the liver from the circulation the portal vein is brought in apposition with the vena cava, the two are sewed together, and a passage opened between them, after which the portal vein is ligated above the anastomosis (forming the so-called Eck fistula). The portal blood then passes directly into the vena cava, and the liver is now supplied only by the hepatic artery. The animals live for a considerable time after the operation, and the urine frequently contains relatively less urea and more ammonia than normal. The results are, however, not nearly so striking as would be expected if the liver were the main seat of urea formation. The experiments have nevertheless brought to light a fact of considerable clinical interest—namely, although the animals may thrive if kept on a diet not containing an excess of flesh, they immediately begin to develop peculiar symptoms, not unlike those of eclampsia or uremia, when they are fed with large amounts of flesh food. Most of the symptoms can be referred to abnormal stimulation of the central nervous system, and examination of the urine has shown a large increase in the excretion of ammonia and a change from the normal acid reaction to an alkaline one.

At one time it was assumed that the toxic symptoms were caused by the presence in the blood of ammonium carbamate, since large quantities of the calcium salt of this substance could be separated from the urine. It is now known, however, that the ammonium carbamate is present only because of the excess of ammonium carbonate, the two salts existing together in solution according to the laws of mass action. That the intoxication is not due to ammonium carbamate does not exclude the pos-

perfused, no urea is formed. The evidence shows that the liver is an important seat of urea formation, but not necessarily that other organs are unable to form it in the intact animal, for there are many *sources of inaccuracy in perfusion experiments.* Even though we exercise the greatest care, we can not hope to maintain the organ in other than a slowly dying condition. It is certainly far removed from the normal state, as is revealed not only by histologic examination, but by the fact that edema almost invariably sets in and the blood vessels become extremely constricted, thus necessitating a gradual increase in the perfusion pressure as the perfusion goes on. Furthermore, the organ being isolated from the nervous system, there can be no control of the relative blood supply of different parts. In the intact animal the circulation is more or less distributed according to the particular needs of the different viscera, and such conditions obviously can not be simulated in a perfusion experiment. Another objection depends on the fact that the well-being of the organs in the intact animal is largely dependent on hormones conveyed to them from other organs. Such hormones are frequently quite labile in nature, and soon disappear from the perfusion fluid.

Notwithstanding these objections, there can be no doubt that many of the functions of an organ are retained much longer than they would be if the organ were not perfused; for example, the contractility of the muscle or the power of forming urea in the liver. Perfusion experiments are of value therefore when they yield positive results. Negative results may indicate either that the organ does not perform the particular function that is being investigated or that it has lost this function as a result of partial death. That a perfused muscle retains its power of contraction does not necessarily indicate that it maintains all of its metabolic functions; neither does the fact that the liver forms urea prove that it is capable of performing its other functions. It is easy to show that the liver dies piecemeal; some functions, such as glycogen-formation, die early, while others, such as urea-formation, remain for a long time intact. *The use of perfusion experiments for the settling of questions of metabolism should therefore always be very carefully controlled and never used as the sole line of evidence on which to base important conclusions.*

(3) Before leaving this subject it may be well to point out that the method which at first sight might appear to be the simplest for settling such questions—namely, *the examination of the inflowing and outflowing blood of different parts or organs*—is not applicable in most cases. This is because of the extremely small changes in concentration which may occur even although large amounts of the particular substance in question

are being absorbed or produced. As we shall see later, this criticism is particularly applicable in the case of sugar. Even during the injection of considerable quantities of sugar into the portal vein, no difference in percentage can be demonstrated between the blood of the two sides of the liver, although we know that sugar is being retained to form glycogen. For the same reasons, differences in the percentage amounts of amino acids or of urea are often difficult to demonstrate in the blood entering and leaving the liver even when we know that large quantities of them are being added to or removed from it.

Clinical.—Since the liver is an important seat of urea formation, the question arises as to whether the relative percentage of urea and ammonia in the urine will become altered by *disease of the liver*. Many observations with this point in view have been undertaken, but it can not be said that the results are very striking. In extreme destruction, such as that produced by phosphorus poisoning, there may indeed be a great increase in the relative amount of ammonia and a decrease in that of urea. The same is true in acute yellow atrophy of the liver, in which disease the nitrogen excreted as ammonia may amount to as much as 70 per cent of that excreted as urea. In milder forms of liver disturbance, however, such as cirrhosis, the figures are much less striking. When an increased ammonia excretion is observed in such cases, we must be cautious in drawing the conclusion that it is due primarily to abolition of the hepatic function. It may just as well be caused by the development of acids in the organism that require the ammonia for their neutralization. It is significant, for example, that considerable quantities of acids are produced in phosphorus poisoning.

Although the urea and ammonia excretions become altered by extensive destruction of liver tissue, it is a remarkable fact that very little if any change occurs *in the amino nitrogen,* either of the urine or of the blood. In experimental necrosis of the liver produced by chloroform or by phosphorus, it is only in the latest stages of the condition and when it is of the very severest type that an amino-acid increase has been found to occur in the blood and urine. The conditions seem to be somewhat different in man, abnormally high amounts of amino nitrogen having been observed in the blood in a considerable proportion of patients with impaired liver function. In very severe cases of diabetes, for example, figures that are distinctly higher than normal have been observed (Van Slyke, etc.). In eclampsia the marked pathologic changes in the liver might be expected to be associated with an upset in the metabolism of amino acids. Losee and Van Slyke[35] have, however, recently shown by the most accurate methods that neither in the blood nor in the urine is any excess of amino acids to be found in this condition, although

in cases of pernicious vomiting of pregnancy, there was a relative increase in the ammonia excretion. We have already seen that this increase did not bear any relationship to the acid-absorbing power of the blood plasma (see page 617).

The importance of the kidneys in removing the urea from the blood is readily seen from the change in the percentage of urea in this fluid after the partial or complete removal of the kidneys. Animals survive nephrectomy for about three days, and during this time urea rapidly accumulates in the blood and begins to make its appearance in the saliva and the intestinal secretions. In man also where the kidneys are extensively diseased, a similar accumulation of urea occurs in the blood, some of the excess being got rid of through the sweat and to a certain extent through the intestine. The importance of encouraging perspiration and a free movement of the bowels in cases of nephritis is thus indicated. It must not be concluded that the accumulation of urea in the organism is the direct cause of the symptoms. Urea itself is comparatively inert, and it is generally believed that other metabolic products with which the urea runs parallel in amount are the toxic agents. Hewlett has found, however, that very large injections of urea do produce symptoms in animals.[34]

CHAPTER LXX

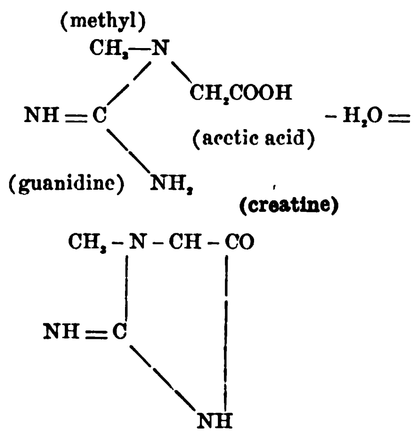
THE METABOLISM OF PROTEIN (Cont'd)

CREATINE AND CREATININE

Creatine and creatinine are very largely products of endogenous metabolism; they are mainly derived from chemical processes occurring in the tissues although some of the creatine and creatinine present in the food may appear as creatine in the urine.

Essential Chemical Facts

Before we proceed further with a discussion of the metabolism of these important substances, it will be necessary to refer briefly to some points in their chemistry. The simpler of the two bodies is creatine, which is methyl-guanidine-acetic acid; creatinine is its anhydrid, being formed from creatine by the removal of a molecule of water, so that the NH_2 groups become joined together in the same way as they do in the formation of peptides from amino acids (page 599). The relationships are illustrated in the following formulas:



(creatine)

(creatinine)

It should be noted that guanidine is closely related to urea $\left(O{=}C\begin{matrix}NH_2\\NH_2\end{matrix}\right)$, and that when creatinine is formed from creatine a ring

formation occurs, giving what may be regarded as an imidazole derivative (see page 604). Creatine is also related to one of the important diamino acids, arginine, since both contain guanidine radicles,

$$\left(NH{=}C\begin{matrix} \diagup NH_2 \\ \diagdown NH_2 \end{matrix}\right)$$

, and to histidine and the purines (see page 634), both of which contain the imidazole ring. The close relationship which creatine bears to urea is illustrated by the fact that urea is formed when creatine is subjected to the action of boiling barium hydrate. When it is oxidized by means of potassium permanganate, urea is also formed, the remainder of the molecule, more or less intact, being split off as methyl-amino-acetic acid

$$\left(CH_2\begin{matrix} \diagup NH\text{-}CH_3 \\ \diagdown COOH \end{matrix}\right)$$

, also known as sarcosine.

The conversion of creatine to creatinine goes on slowly in aqueous solutions, but is much accelerated by heating with acid. Heated in an autoclave at a temperature of 117° C. for thirty minutes, with half normal hydrochloric acid, the creatine goes over almost quantitatively into creatinine. It will be noted that the creatinine ring is partly oxidized. This renders it unstable, so that creatinine in the presence of alkalies has the power of reducing metallic oxides. Like glucose it can reduce alkaline solutions of copper, silver and mercuric salts; it also reduces picric acid in weakly alkaline solution to picramic acid, which, being red, furnishes us with a solution the strength of which can be estimated colorimetrically.

Quantitative Estimation.—Although the presence of creatinine in the urine has been known for many years, there being from 1 to 2 grams of it in the twenty-four-hour urine, little progress was made in the study of its metabolism because of the absence of a reliable method for its estimation. The elaboration by Folin of a colorimetric quantitative method for *creatinine,* depending on the reduction of picric acid, has furnished the starting point for the modern work which has been done. To estimate the *creatine* by this method, it is usual to proceed as follows: The creatinine content is first of all determined, another portion of urine being then heated with acid in the autoclave until all of its creatine has been converted into creatinine. A second determination of creatinine is then made, and the difference between the two is calculated as creatine.

It should be pointed out that, since the creatine is estimated by an indirect method, there are considerable chances for inaccuracy. Indeed, it has been shown that errors may have been incurred in some of the recent work on account of the fact that when acetoacetic acid is present in the urine it prevents the creatinine from developing its full reducing power on picric acid in the cold, so that when subsequently the urine is heated with acid for the purpose of converting the creatine into creatinine, the destruction of acetoacetic acid allows the reducing power of the creatinine to develop to full intensity. It is obvious that this would make it appear as if creatine had been converted into creatinine. It is particularly in the urine of diabetic patients, in which acetoacetic acid is present that mistakes are likely to be made.

Metabolism

When we come to consider the metabolism of creatine and creatinine, we find that there are remarkably few facts definitely known concerning it. The average amount excreted daily, expressed as the number of milligrams of creatinine in twenty-four hours per kilogram body weight, is known as the *creatinine coefficient* (Shaffer).[36] For a lean person this is about 25 mg.; for a corpulent person, about 20 mg., the difference indicating that muscle mass, and not body weight, is the important factor determining the coefficient. Further evidence that this relationship exists is furnished by the fact that in the muscular atrophies creatine excretion is distinctly below normal. It must be the mass of the muscles rather than their activities that is the determining factor, for the creatine excretion does not become increased by muscular exercise.

Influence of Food, Age, and Sex.—Although creatine and creatinine are endogenous metabolites, it must be remembered that, under ordinary dietetic conditions, a part of each is derived from these substances present in the food. It is important therefore to consider the conditions under which the creatine and creatinine in the food *appear in the urine.* Regarding creatinine, it is pretty well established that practically all that is taken with the food reappears as creatinine in the urine. Shaffer has, for example, succeeded in recovering 76 per cent of ingested creatinine in the urine excreted during twenty-one hours following the ingestion of 0.7 gm. creatinine.

The conditions for the excretion of creatine are more complex. It is present in the urine of children in considerable amount, but in that of adults only as traces. In the first years of life the creatine in boys' urine may amount to one-half of the total creatine and creatinine, but it becomes gradually less and practically disappears at about seven

years of age. Girls, on the other hand, continue to excrete creatine until about puberty, after which, although ordinarily absent, it reappears in the urine at each monthly sexual cycle, and is present during pregnancy and for some days after delivery. Feeding creatine to children causes it to appear in the urine, accompanied usually by a slight increase in the creatinine. The same results can be observed in women during the monthly periods, when as much as 0.1 gm. may be present, and during pregnancy. Creatine is also present in the urine of most if not all of the other mammalia. Some of these facts are shown in the following table:

	AGE	CREATININE-N	CREATINE-N EXCRETED IN 24-HR. URINE
Boys	2	0.025	0.023
	3	0.057	0.022
	5	0.112	0.025
	8	0.163	0.0
	11	0.157	0.0
	15	0.378	0.0
Girls	5	0.069	0.005
	6	0.032	0.003
	7	0.157	0.066
	10	0.147	0.020
	12	0.201	0.011

(From Mathews.)

When creatine is given to an animal that has been kept in a starved condition, most of it seems to disappear. It can not be recovered in the urine either as creatine or as any other nitrogenous metabolite. It seems to functionate more as a food than as a useless substance. The possibility that some of it can be destroyed by the intestinal bacteria being admitted, there is nevertheless some justification for the view that the creatine finds a useful function in the anabolic process of the muscles.

Influence of Complete and Partial Starvation.—Although, as we have seen, the creatinine excretion remains constant when the amount of protein in the diet is greatly reduced, yet it does not remain constant during complete fasting or when carbohydrates are entirely withheld from the diet. In fasting it has been found that creatine appears in place of the creatinine which has disappeared, so that if both creatine and creatinine are determined, very little if any diminution will be found to have occurred. Fasting, therefore, causes the adult creatine and creatinine metabolism to become like the juvenile metabolism. As pointed out by Mathews, it would be interesting in the light of this observation to see whether other substances, passed in the urine of young animals but absent in that of the adult, would reappear in the urine when the animals were made to fast. In the case of man, for instance, allantoin would be worth investigating in this regard (page 641).

A similar replacement of some of the creatinine by creatine appears when carbohydrate is entirely withheld from the diet, or in diabetic animals, either in the disease diabetes mellitus in man or in the experimental condition induced in animals by giving phlorhizin. Unfortunately, in a considerable part of the work that has been done on this phase of the subject a method of estimation was employed which did not take sufficiently into account the influence of acetoacetic acid on the creatine estimation; but even after allowing for this possible source of error, there can be no doubt that creatine appears in the urine when carbohydrates are improperly metabolized. If carbohydrates are given to a starving animal, for example, the creatine is replaced in its urine by creatinine, although this will not occur when either protein or fat is fed. The general conclusion which may be drawn from these observations is that carbohydrates in some way are required for the proper conversion of creatine into creatinine in the animal body (Cathcart)[37].

Origin of Creatine and Creatinine

Notwithstanding the amount of excellent work that has recently been done on the metabolism of creatine and creatinine, we know very little indeed regarding the origin of these bodies in the animal organism. It would be profitless to discuss this problem to any great extent, but a few of the most important facts so far established may be of interest and of value. The first step in attacking such a problem is to compare the amounts present in the various organs and tissues, in the blood, and in the excreta. Of the approximately 120 grams of creatine and creatinine in the body of an average adult, a very large proportion is in the muscles, the voluntary muscles containing the largest percentage, the heart containing a medium percentage, and the involuntary (intestinal) muscles containing relatively a small amount (Myers and Fine)[38]. Next to the skeletal muscles, and containing more than the involuntary muscles, come the testis and brain. The liver, pancreas, thyroid, kidneys, spleen, etc., contain traces, the smallest amount of all being found in the blood.

In all these places by far the greatest proportion of the total creatine-creatinine exists as creatine, which is exactly the reverse of the condition obtaining in the urine of adults, where practically all is excreted as creatinine. The close chemical relationship between creatine and creatinine, considered along with the above facts regarding their quantitative distribution in the body, indicates that the creatinine of the urine is derived from the creatine of the tissues. The question is, *How does the creatine come to be converted into creatinine?* Such a transformation is

certain aromatic alcohols are given, when it is used for the purpose of detoxicating them. The resulting glycuronates have reducing powers and may be confused with glucose when present in large amount. Glycuronates may be distinguished from glucose in the urine (1) because they are levorotatory, and (2) because they do not ferment. The free acid itself, however, is dextrorotatory.

CHAPTER LXXII

URIC ACID AND THE PURINE BODIES

Introductory.—The participation by highly trained organic chemists in the investigation of biochemical problems has brought our knowledge of the history of the purine substances in the animal body from a state of chaos and guesswork to one of system and scientific accuracy. The peculiar solubility reactions of uric acid and its salts and the discovery of urates in gouty deposits served to make uric acid metabolism one of the earliest research problems in both the medical clinic and the biochemical laboratory, but the earlier results were practically valueless, partly because they were inaccurate and partly because their interpretation was impossible in the absence of even the most elementary facts concerning the chemistry of uric acid.

Before any real progress was possible, a clean sweep had to be made of all the old speculations and hypotheses, such as that dignified by the high-sounding name of "uric-acid diathesis," and a foundation of accurate chemical knowledge established. This foundation is now wonderfully complete, and a superstructure of biochemical fact is already beginning to grow upon it. In the present chapter we shall examine some of the most important contributions that have made this progress possible.

As in the study of any other problem of metabolism, we must, however, make ourselves familiar with the main facts concerning the chemistry of the purine bodies and of the tissue constituents into the composition of which they enter, before proceeding to the more strictly biological aspect of the subject.

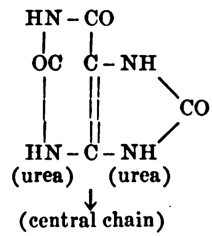
The Chemical Nature of the Purines

By an examination of the empiric formulas of the purines of biochemical interest, it will be observed that they are all derivatives of a substance purine, which although in itself of no importance is interesting, since it serves as the basic substance from which the others are derived. The list is as follows:

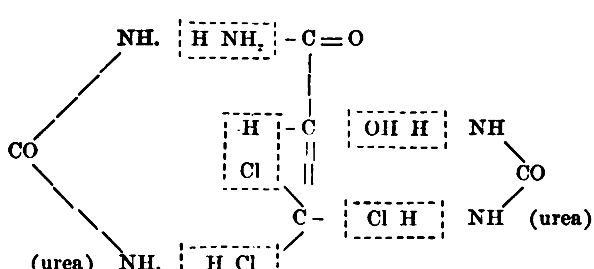
Purine . .	$C_5H_4N_4$		
Hypoxanthine .	$C_5H_4N_4O$	Monoxypurine	Purine bases.
Adenine . .	$C_5H_3N_4.NH_2$	Amino-purine	
Xanthine . .	$C_5H_4N_4O_2$	Dioxypurine	
Guanine . .	$C_5H_3N_4O.NH_2$	Amino-oxypurine	
Uric acid . .	$C_5H_4N_4O_3$	Trioxypurine	

The first oxidation product of purine is hypoxanthine, which has long been known as a constituent of meat extract. Adenine, the amino derivative of hypoxanthine, occurs in combination with other substances in the nuclear material. The second oxidation product is xanthine and its amino derivative, guanine. They occur in the same places as hypoxanthine and adenine. The highest oxidation product of all is the well-known urinary constituent, uric acid, which may therefore be chemically designated as trioxypurine. In addition to the purines of animal origin, there are also certain ones of vegetable origin—the methyl purines, which exist as the alkaloids of tea and coffee—namely, caffeine, theobromine, and theine.

To understand the chemical structure of this group of substances, it is perhaps simplest to start with that of uric acid. This consists essentially of two urea molecules linked together by a central chain of three carbon atoms, as will be evident from the accompanying structural formula:



This structure can be shown by methods both of decomposition and of synthesis. When uric acid is decomposed by oxidizing it with nitric acid, it yields urea and a residue called alloxan; or it can be synthesized from urea and trichlorlactamide, a derivative of lactic acid, which it will be remembered contains three carbon atoms. The changes involved in this synthesis will be made clear by examination of the accompanying structural formula, in which the manner of production of the by-products of the reaction (NH_3, H_2O and HCl) are shown by dotted lines:



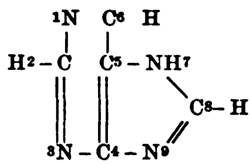
(trichlorlactamide)

By milder oxidation by means of potassium permanganate in the cold, uric acid becomes quantitatively converted to allantoin:

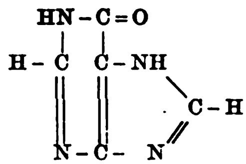
$$\underset{\text{(uric acid)}}{C_5H_4N_4O_3} + H_2O + O = \underset{\text{(allantoin)}}{C_4H_6N_4O_3} + CO_2.$$

The importance of this transformation lies in the fact that in most animals, man and the higher apes being exceptions, uric acid is thus decomposed in the animal body. The structural formulas for the other purine bodies in relationship with those of purine and uric acid are given below.

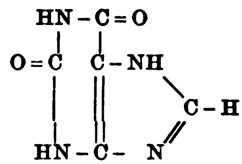
Purine itself has the following structural formula:



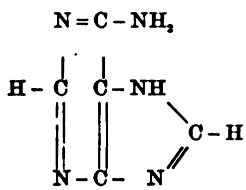
(For convenience of description the atoms in purine are numbered as shown.)



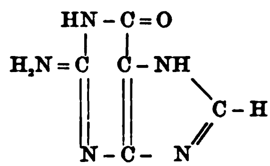
(hypoxanthine) (6-oxypurine)



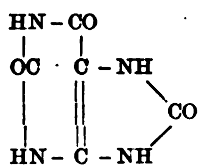
(xanthine) (2-6-oxypurine)



(adenine) (6-amino-purine)

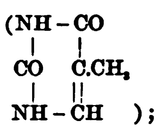
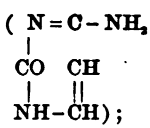
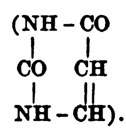


(guanine) (2-amino-6-oxypurine)



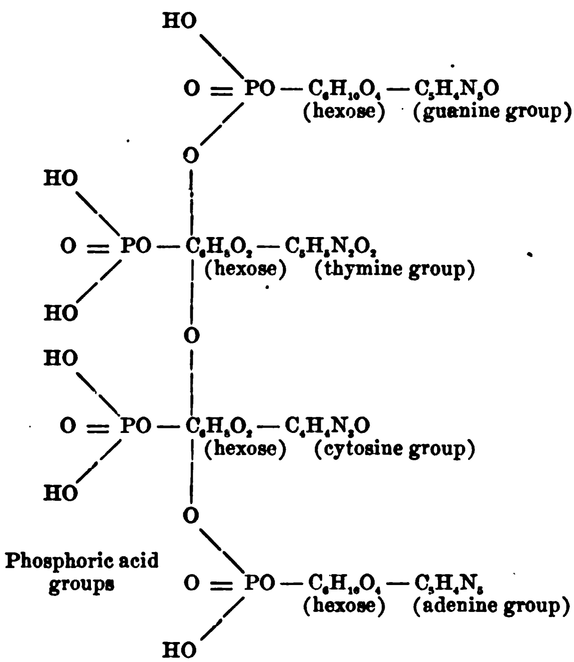
Uric acid (2-6-8-trioxypurine)

The substances with which the purine bases are most closely related are the *pyrimidine bases.* Three of these are known:

thymine ; cytosine ; and uracil .

From an examination of the structural formulas, it will be seen that they are more or less related to purine (having one of the urea radicles omitted), although it can scarcely be doubted that they exist as separate constituents of the nucleic acid group in the animal body, and are not derived from purine. They are primary products.

The Chemical Nature of the Substances in Which Purine and Pyrimidine Bases Exist in the Animal Body.—In general it may be said that the amino purines—adenine and guanine—together with the pyrimidine bases—thymine and cytosine—occur combined with phosphoric acid and a carbohydrate in the various *nucleic acids,* each of which is again combined with some simple protein to form nuclein, the essential constituent of the chromatin of the nucleus. One of the oxypurines, hypoxanthine, may also exist combined with phosphoric acid and carbohydrate to form a substance present in muscle and known as inosinic acid. The general scheme of construction of a nucleic acid of animal origin is illustrated in the following formula suggested by Levene and Jacobs:[39]



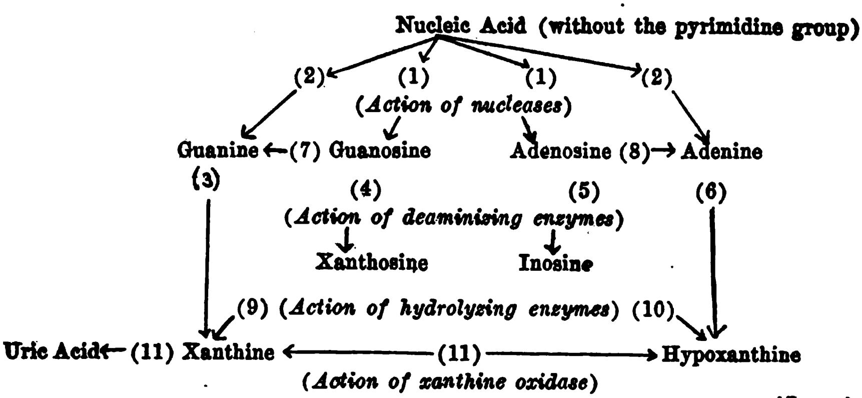
According to this formula nucleic acid may be considered as a compound of polyphosphoric acid, containing carbohydrate groups, which serve to link the phosphoric acid molecules to those of purine or pyrimidine. In nucleic acids of animal origin, such as the example given above, the carbohydrate is a hexose, (i.e., contains 6 C-atoms), whereas

in those of plants (e. g., yeast), it is a pentose (5 C-atoms). It has been found necessary to introduce some terms to designate the different parts of the nucleic acid molecule; thus, the whole molecule is called a *tetranucleotide,* each mononucleotide molecule of which is composed of a phosphoric acid molecule *plus* a *nucleoside,* which again is composed of a purine or pyrimidine nucleus attached to pentose or hexose. The nucleoside is so named because it is similar in structure to a glucoside.

Apart from differences in the carbohydrate group, it appears that there is a close similarity in the structures of nucleic acids from different cells. This would indicate a common function for them all, which may be either of a skeletal or of a physiologic nature; that is, nucleic acid may have to do with the sustentacular material that builds the nucleus, or it may have to do with some physiologic function common to all cells, such as irritability, or growth, or respiration. If nucleic acid is merely a sustentacular material, then the study of the behavior of chromosomes and chromatine in cells can not have the significance that it would have were nucleic acid concerned in the more vital activities of the nucleus. All the so-called nuclear stains owe their specific staining properties to the fact that they are of a basic nature and combine with nucleic acid. Until we know more definitely what the exact function of nucleic acid may be, it is unwise to place too much weight on the behavior of the chromosomes in cytologic researches.

The History of Nucleic Acid in the Animal Body.—We shall first of all study the manner in which nucleic acid may be broken down. As is to be expected from its complex structure, various types of enzymes are concerned in this process. The first to act are known as the *nucleases.* They split the tetranucleotide molecule into two dinucleotides, which immediately afterward split further into mononucleotides. Four nucleotides, two of purine and two of pyrimidine, are thus formed from each molecule of nucleic acid. Each nucleotide molecule may now undergo decomposition in one of two ways: (1) either by the splitting off of phosphoric acid, leaving a nucleoside (guanosine or adenosine), or (2) by the splitting off of both phosphoric acid and carbohydrate, leaving free purine bases. *Nucleases* have been found which specifically effect either of these decompositions, and they have been called phosphonucleases* (1), and purine-nucleases (2), respectively. In the decomposition of nucleic acid all of the four purine compounds—guanine, guanosine, adenosine and adenine—may be formed. This is illustrated in the accompanying schema, in which the nucleic acid is represented as a purine nucleotide:

*The numbers refer to the enzymes indicated in the schema.



(Jones.)

The next step in the disintegration process is that the amino group is removed and the corresponding oxypurine is produced. To bring this about, there exists a specific *deaminizing enzyme* for each of the above amino compounds, and each enzyme is named according to the exact amino purine upon which it acts; thus, guanase (3), guanosine-deaminase (4), adenosine-deaminase (5), and adenase (6) have all been identified. The free base may then be split off from the nucleosides by specific *hydrolyzing enzymes* (7) (8) (9) (10).

The joint action of these enzymes leads to the formation of oxypurines, xanthine and hypoxanthine, which are oxidized to uric acid by *xanthine-oxidase* (11).

In man and the anthropoid apes uric acid is the end product of the above changes, but in other mammals most of the uric acid is further oxidized into allantoine. It has also been found, except in man and the chimpanzee, that extracts of organs such as the liver, are capable of decomposing uric acid into allantoine. The identification of these specific enzymes is sought by a determination of the free amino-purine bases and the phosphoric acid produced by allowing an aqueous extract of the tissue in question to act on nucleic acid (of yeast)* at body temperature. Another portion of the digested mixture is then hydrolyzed by means of boiling sulphuric acid and the constituents again determined. From the results it is often possible to draw conclusions as to the exact nature of the enzymes present.

The most remarkable outcome of this work has been to show that *the distribution of the enzymes is not the same in the tissues and organs of different animals.* Very briefly, some of the most important results that have so far been obtained are as follows: Gastric and pancreatic juices do not contain a trace of any of the enzymes. Intestinal juice,

*Yeast nucleic acid is used because it is less resistant to disintegration than thymic nucleic acid.

on the other hand, contains a nuclease capable of splitting the polynucleotides into mononucleotides. The two pyrimidine nucleotides split off do not undergo further change, but the purine nucleotides are converted into nucleosides (the enzyme being designated "nucleotidase"). Extract of the intestinal mucosa, besides having the same action as the intestinal juice, can also decompose the purine, but not the pyrimidine nucleosides, into carbohydrate and purine groups (specific action of "nucleosidase"). A similar action is produced by extracts of kidney, heart muscle, and liver. Blood serum, hemolyzed blood, and extract of pancreas, on the other hand, are capable only of carrying the decomposition as far as the mononucleotides.

Regarding the other enzymes mentioned in the above list, it is important to note that they appear at different stages in embryonic development, and that their distribution varies considerably in different species of adult animal, the spleen, liver, thymus, and pancreas containing them most abundantly. The distribution of enzymes in the organs of the monkey resembles that in the lower animals considerably more than it does that in man. Some remarkable facts have come to light regarding guanase and adenase, particularly that guanase is deficient in the organs of the pig, in the urine of which animal it has also been found that the purine bases are in excess of the uric acid. This absence of guanase no doubt accounts for the fact that deposits of guanine may occur in the muscles, and that these may be so large as to constitute the condition known as guanine gout found in this animal. Adenase, on the other hand, is absent from the organs of the rat, which again corresponds with the fact that, when adenine is injected subcutaneously into these animals, it undergoes oxidation without the removal of its amino group. In the human organism, adenase appears to be absent from all of the organs, whereas guanase is present in the kidney, lung and liver, but not in the pancreas or spleen. Xanthine-oxidase exists only in the liver.

The distribution of *uricase* is perhaps the most interesting. It is present in most of the lower animals. On account of its presence extracts of the liver, spleen, etc., in all breeds of dogs, with the exception of Dalmatians, rapidly destroy uric acid; and practically no uric acid when injected subcutaneously can be recovered unchanged in the urine, but appears as allantoine. Uricase, however, is absent in man. This has been demonstrated by finding (1) that when uric acid is injected subcutaneously, nearly all of it appears in the urine, and (2) that uric acid is not destroyed when extracts of the organs are incubated at body temperature with uric acid or its precursors. It must of course be kept in mind that, although the uric acid is thus shown not to be destroyed *in vitro,* it may nevertheless be destroyed in the living animal.

The importance of the above described results rests in the fact that from them we may hope to be able, ultimately, to state exactly in what organs and tissues the intermediary metabolic processes concerned in nucleic acid metabolism occur. The work at the present time is of special significance, since it represents one type of evidence which we must have before we can trace exactly every step in the metabolism of any other biochemical substance.

The absence of uricase from the tissues of man places him in a unique position with regard to the metabolism of nucleic acid, and renders the investigation of the problem particularly difficult, since animal experimentation is useless. Recently, however, S. R. Benedict has discovered that the Dalmatian breed of dog—also known as the carriage dog, and having a spotted or mottled skin—has a purine metabolism like that of man.[4] When fed on food containing no purine substances, he excretes large quantities of uric acid, and when the latter substance is injected subcutaneously, it is eliminated quantitatively as such in the urine. We shall see later how experiments on this animal have been made use of in the investigation of problems of purine metabolism as applied to man. In all other animals most of the uric acid is oxidized to allantoine before being excreted. The degree to which this occurs varies between 79 and 98 per cent of the uric acid in different species. This has been called the uricolytic index (Hunter and Givens).

The Balance between Intake and Output of Purine Substances under Various Physiologic and Pathologic Conditions.—The main purine excretory product in man is uric acid, but there is also a certain amount of purine bases. The presence of uric acid has attracted attention for a great many decades in medical investigation, because of the relative ease with which it can approximately be determined quantitatively, and because of the well-known fact that it may be responsible for certain diseases, such as gout, when it accumulates in the tissues in an insoluble form. On a diet containing meat, or more particularly on one containing glandular substances, the total daily excretion of uric acid is very considerably greater than when the diet contains no such food stuffs. The conclusion which Burian and Schur[43] drew from this observation is that purine must be partly of *exogenous* and partly of *endogenous* origin. In other words, some of it is derived more or less directly from performed purine substances in the food, and the remainder from the purine constituents of the animal's own tissues.

Endogenous Purines.—It was thought that a definite proportion of each of the administered purines could be invariably recovered from the urine. Although this has not been found to be exactly true, there is nevertheless a certain constancy in the proportion of administered

purine that is excreted. Thus, Mendel and Lyman have found recently that about 60 per cent of injected hypoxanthine, 50 per cent of xanthine, 19-30 per cent of guanosine, and 30-37 per cent of adenine were eliminated as uric acid. When combined purines—i. e., nuclear material—are given, only a small proportion of the purine reappears as uric acid in the urine. There is, therefore, a general parallelism between the purine content of the food and that of the urine, which indicates that purine-rich food should be eliminated from the diet of patients who are suffering from deposition of insoluble urate in the tissues, as in gout. The fate of the purine that disappears in the body is unknown; some of it may be decomposed in the intestine, but why so much of the remainder, after absorption by the blood, should disappear is a mystery, since no uricase can be discovered in any of the organs or tissues. The destroyed purines can not be shown to influence any of the other well-known nitrogenous metabolites of the urine.

The following table of experiments by Taylor and Rose[45] may serve to illustrate these points. The subject was placed on a purine-free diet consisting of milk, eggs, starch and sugar, for three days. After this period a part of the total nitrogen (3 grams) was supplied as sweetbreads—thymus gland, etc.—containing a high percentage (0.482) of purine nitrogen; for another period of four days still more of the nitrogen (6 grams) was replaced by sweetbread nitrogen; and this was followed by a final period in which the original diet of milk, etc., without purine substances, was restored. The following table gives the results:

	1st Period Purine-free Diet	2nd Period	3rd Period	4th Period Purine-free Diet
Total urinary N	8.9	8.7	9.1	8.8
Urea N and NH_3	7.3	7.1	7.1	7.05
Creatinine	0.58	0.55	0.56	0.47
Purine N (total)	0.11	0.17	0.26	0.10
Uric acid N	0.09	0.14	0.24	0.07
Remainder N	0.91	0.88	1.18	1.18

The increase of uric acid accounted for less than half of the purine nitrogen ingested. This appeared as uric acid, the excretion of purine bases being practically unchanged.

CHAPTER LXXIII

URIC ACID AND THE PURINE BODIES (Cont'd)

SOURCE OF ENDOGENOUS PURINES

Even after the entire elimination of all purine substances from the food in the case of man, purine continues to be excreted in the urine as uric acid. This, as above remarked, is called endogenous excretion. At first it was thought by Burian and Schur that the total nitrogen of the purine-free diet could be considerably varied without causing any alteration in the amount of the endogenous purine excretion, but a repetition of the work has shown that, when these changes are of considerable magnitude, the endogenous moiety does not remain constant. This has already been demonstrated in the table on Folin's results (see page 614), and is still better illustrated in the accompanying table, which shows the excretion of uric acid and coincidently of urea from hour to hour in the urine after taking food which is free from nuclein or purine substances. After a fast of six hours, a diet consisting of bread and potatoes was taken at 1:30, and the urea and uric acid measured in the urine each hour thereafter.*

TIME	UREA GM.	URIC ACID MG.	AMOUNT OF URINE C.C.
10-11	1.07	26	175
11-12	1.13	27	118
12-1 P.M.	1.07	24	164
1-2 (meal)	0.64	21	60
2-3	1.12	22	43
3-4	1.16	38	41
4-5	0.84	40	53
5-6	1.16	56	59
6-7	1.20	39	56
7-8	1.37	30	95
8-9	1.47	33	183
9-10	1.33	24	155
10-11	1.33	23	180

(Hopkins and Hope.)[46]

A postprandial increase of endogenous purine excretion is very distinct, and it indicates that during the process of assimilation something must be occurring in the organism which entails the production of purine

*These investigations should be repeated, since there is some question as to whether the method of analysis employed (Folin-Shaffer) is suitable for determining *hourly* uric-acid excretion.

from the organism itself. As to what this may be, it is impossible to say. It may be associated with the work of the gastric and intestinal glands, which recalls the interesting suggestion, originally made by Horbaczewski, that ingested substances increase the excretion of uric acid by causing a leucocytosis, the purine being derived from the nucleic acid set free when the leucocytes become broken down. That this is not the correct explanation, however, is indicated by the fact that ingested substances that give rise to an increased number of leucocytes affect the excretion of uric acid *during* the period the leucocytes are present in the blood, and not *after* they have disappeared, which would have to be the case were the uric acid a product of purine substances liberated by their breakdown. This would indicate that the purine substance is a metabolic product of the living leucocytes and not a breakdown product of those that are dead. It should be noted that the increase in the postprandial uric-acid excretion occurs earlier than that of urea.

The most pressing question concerns *the origin of the endogenous purines.* Uric acid is the purine with which we are most concerned in the case of man, and chemistry shows us that it may be produced either by the oxidation of the lower purines—namely, of those which are the constituent parts of the nucleic-acid molecule—or by a synthesis of two urea molecules with a carbon residue containing three carbon atoms. There are consequently two sources from which the endogenous purine excretion in man may be derived : (1) synthesis of two urea molecules, and (2) oxidation of the lower purines.

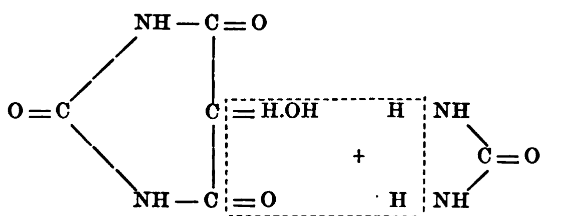
We will consider first **the possibility of synthesis.** In birds and reptiles practically all the nitrogen is excreted in the form of uric acid, and it is easy to show that this has been produced in the organism by the synthesis of urea with carbon-rich residues, occurring mainly in the liver. Minkowski found that by removing the liver from geese, which is a comparatively simple operation on account of an anastomotic vein between the portal and the renal veins, the uric acid in the urine became very markedly decreased and ammonium lactate took its place (page 618). Since we know that ammonium in the animal body is ordinarily converted into urea, we may conclude from this observation that something has occurred to prevent the synthesis of urea into uric acid. In confirmation of this conclusion it was subsequently found that, if ammonium lactate was added to the blood perfused through the isolated liver of the goose, uric acid was produced in the perfusion fluid.* Furthermore, when birds and reptiles are fed with ammonium salts or with the degradation products of protein, there is an increase in the ex-

*The reason for the formation of this relatively insoluble metabolite in place of the soluble urea is connected in some way with the fact that birds and reptiles do not take such large quantities of water with their food as other animals.

cretion of uric acid instead of urea. Everything which in a mammal tends to cause an increase in urea excretion causes in birds and reptiles a similar increase in the excretion of uric acid.

In the early days of research in the uric-acid problem, not inconsiderable mistakes were made on account of failure to recognize the essential difference in the metabolism of uric acid in birds and mammals, and the tendency for some time after the exact state of affairs was discovered was to consider that in mammals none of this synthetic process occurs. The latter view, however, is surely incorrect, for a *certain amount not only of uric acid itself but of the lower purine bodies can be produced by synthesis in the mammalian body.* Thus, Ascoli and Izar[47] discovered that uric acid could be made either to disappear or to be formed when a minced preparation of liver was incubated, depending upon whether oxygen or carbon dioxide was bubbled through it. With oxygen uric acid disappeared, whereas with carbon dioxide uric acid accumulated, indicating that in the presence of this gas the destroyed uric acid became reformed from the disintegration products of the oxygenation process. As similar results were obtained from the livers of birds, it is clear that no essential difference exists between the purine metabolic processes occurring in the livers of birds and of mammals. The difference is a quantitative not a qualitative one.

Regarding the chemical nature of the product into which uric acid is broken down and from which it may be resynthesized, it has been possible so far to identify but one substance—namely, dialuric acid. This is a perplexing result, for from all other investigations it would appear that in mammals, with the exception of man and the anthropoid apes, uricase splits uric acid into allantoine (see page 640), which substance, however, when added to liver extract did not cause any uric acid to be formed; nor did any of the other known decomposition products of uric acid have such a result. The chemical reaction involved in the production of uric acid from dialuric acid and urea is indicated as follows:



(dialuric acid) (urea)

The synthesis of uric acid is brought about by the combined action of a thermolabile enzyme in the blood and a thermostable body in the tissues. An aqueous extract of blood-free liver of the dog can destroy

uric acid only in the presence of oxygen; it can not reform it even in the presence of carbon dioxide. On the other hand, blood serum can not reform uric acid, whereas a mixture of the bloodless liver extract and blood serum produces uric acid readily under suitable conditions. Boiling of the liver extract does not affect the result, but boiling of the blood serum renders it incapable of exerting its joint action with the bloodless liver extract.

These experiments with dog's liver serve only as circumstantial evidence that uric-acid synthesis occurs in mammals as well as in birds. More *direct proof that purine synthesis occurs in mammals* is as follows: (1) It was discovered long ago by Miescher that salmon, after leaving the sea to ascend the rivers, have a well-developed muscular system, but that in the upper reaches of the stream the muscular system becomes considerably atrophied and the testes enormously developed. As the fish takes no food during the migration, there must be conversion of the protein of the muscles into the cellular tissue of the sexual glands, and nucleic acid must be produced. (2) A hen's egg before its incubation contains practically no nucleic acid, whereas after development has well started nucleic acid increases by leaps and bounds. Similarly the eggs of insects increase in purine content very markedly as development proceeds. (3) Milk contains practically no purine derivative, and yet when it is fed to young growing animals, the organs lay on purine substances abundantly. In general, indeed, it may be said that the combined purine increase is in proportion to the increase in body weight on the milk diet. (4) In Osborne and Mendel's experiments already alluded to, it has been shown that adequate growth depends primarily on the nature of the protein building stones, and not upon the purine content of the food. (5) An objection might be raised to these results on the score that they do not apply to the adult mammal. Investigation of the problem has hitherto been seriously impeded by the fact that no ordinary laboratory animals were known in which uric acid is excreted in the urine. The discovery that this occurs in the Dalmatian dog has, however, made it possible for S. R. Benedict[41] to show, not only that after increasing the amount of nonpurine food there was a very distinct increase in the uric-acid excretion, but also that when the animal was kept for a year on such foods there was excreted a total amount of uric acid at least ten times greater than could have come from the traces unavoidably included in the food.

Regarding the chemical nature of the substance from which the purine is synthesized, we know at present practically nothing. No doubt some of the protein building stones functionate in this capacity, pyrimidine being probably the product that is first formed. Thus, pyrimidine may

be produced as a result of the combination of amino-malonic acid with urea, the amino-malonic acid being produced by condensation of hydrocyanic-acid molecules:

$$
\begin{array}{cccc}
3\,HCN \rightarrow & H_2N-CH(CN)_2 + & CO(NH_2)_2 \rightarrow & \begin{array}{l} NH-CO \\ | \qquad | \\ CO \quad CNH_2 \\ | \qquad \| \\ NH-CNH_2 \end{array} \\
\text{(hydrocyanic acid)} & \text{(amino-malonic nitrile)} & \text{(urea)} & \text{(oxy-diamino-pyrimidine)}
\end{array}
$$

Another possible source of pyrimidine is the oxidation of arginine to guanidine-propionic acid, which then condenses to form amino pyrimidine.

Purine synthesis undoubtedly occurs in the mammalian body, but it is difficult to recognize in metabolism investigations because it is a slow, continuous process. The probability of its occurrence, however, is indicated by such results as those described on page 614, in which increase in purine excretion is observed after varying the intake of food, even when this is itself entirely free from purine substances. Whether or not changes in the activity of purine synthesis occur in conditions of disease is a question which awaits investigation.

The Influence of Various Physiologic Conditions, of Drugs, and of Disease on the Endogenous Uric-acid Excretion.—*Muscular exercise* was thought by Burian to cause an increased excretion of uric acid, from which he drew the conclusion that the hypoxanthine present in comparatively large amount in muscular extract, or its precursor, inosinic acid, must be an important source of endogenous uric acid. Other observers (Leathes, etc.) have found that *strenuous* exercise causes a distinct increase in uric-acid excretion, which, however, is much less marked on repetition of the same kind of exercise on the next day. If some new kind of muscular work is performed, another increase in uric acid will result. There are still other investigators who deny that muscular work has any influence on uric-acid excretion.

It has been observed by several investigators that the endogenous purine excretion is distinctly *higher during the waking hours than during sleep*. This can not be shown to depend on variations in the urinary function, and since it is decidedly doubtful whether ordinary muscular activity has any influence, the diurnal variation is most difficult to account for. The endogenous excretion in man is not the same for *different individuals*, even when calculated for the same body weight; it varies between 0.12 and 0.20 per cent purine nitrogen in an adult man. It remains remarkably constant for a given individual from time to time, being unaffected by moderate degrees of variation in the amount

of food taken provided this be purine-free; when, however, the amounts are extremely variable, changes are produced (see page 614).

In disease, *fever* causes an increased excretion. This has been most clearly shown by Leathes, who took a large enough dose of antityphoid serum to produce a distinct degree of fever (103° F.), and found that an increase in uric-acid excretion occurred. That increased combustion processes occurring in the tissues were responsible for the uric acid, was shown by the same author, who caused a similar increase by subjecting himself to cold baths for a considerable period of time. The increased loss of heat thus induced stimulated the combustion processes in the body so as to maintain the body temperature, and as a result there was an increase in uric-acid excretion. It has long been known that an excessive amount of uric acid is excreted in *leucocythemia*. The nuclein of disintegrated leucocytes is commonly held responsible for the increase. Naturally, much work has been done on the endogenous and exogenous purine excretion in *gout*. No very striking anomalies of excretion have, however, been brought to light, except perhaps that after the ingestion of purine-rich foodstuffs it takes longer for the resulting exogenous excretion to develop and pass away.

Certain *drugs* affect the excretion of uric acid. Salicylic acid is said to cause an increased excretion, and citrates certainly have this effect. In both cases the increase is followed by a compensatory fall, which indicates that these drugs act by facilitating the excretion rather than by influencing the metabolic processes that are the source of the uric acid. The effect of caffeine has been very carefully investigated. Given to the Dalmatian dog, referred to above, S. R. Benedict found that a small dose caused a slight decrease, but that a larger dose had practically no effect, although there was a notable retention of nitrogen. On man, however, different results were secured, for it was found that when 1 gram of caffeine was given daily for several days, a slight but definite progressive increase in the endogenous uric-acid excretion occurred, and it lasted for 10 days after the caffeine administration was discontinued. Liberal allowance of this alkaloid may, therefore, not be quite so innocuous as it is assumed to be.

Uric Acid of Blood.—In all of the investigations considered above, the behavior of uric acid is judged from the amount of it excreted in the urine. Valuable though such results must be, their interpretation is always difficult, since two factors that are quite independent of each other have to be kept in mind—namely, the production of the uric acid in the organs and tissues and its excretion by the kidneys. In connection with the latter factor, we must also consider the method of transportation of uric acid by the blood from its place of production (or absorp-

diabetes. For this purpose the patient should first of all be instructed to follow his usual diet, so that, by examination of the amount of sugar excreted in the urine, an opinion may be formed of the severity of the case. The diet should then be changed so as to consist of a part that contains no carbohydrates and another composed entirely of starchy food. The former is made up of eggs, fish, green vegetables, fat, etc., and the latter, to start with, should consist of 100 grams of bread, distributed between the two main meals of the day, one of which is breakfast. This diet should be continued until the glycosuria either disappears or attains a constant level. If it disappears, the case is classified as a *mild* one of diabetes, and the daily allowance of bread may be increased, by 50 grams a day, until the sugar again makes its appearance in the urine, indicating that the assimilation limit has been reached. For therapeutic purposes, the patient should now be instructed to take about three fourths of this amount of carbohydrate in his daily rations, and he should be supplied with explicit instructions in the shape of diet tables as to what variety and quantities of the various carbohydrate materials his food may contain. His urine should be examined at frequent intervals—once a week—and he should be instructed as to the nature of his disease and the importance of his remaining aglycosuric. By further treatment such so-called latent cases of diabetes may be kept in perfect health for many years.

When, on the other hand, the glycosuria exists with 100 grams of bread in the daily ration, this must be reduced to 50 grams, and if after some days the first reduction does not suffice to render the urine free from sugar, carbohydrates must be withheld entirely from the diet. If the glycosuria does not now disappear, the case is to be considered *severe,* and it may be necessary to undertake the starvation treatment, which has recently been developed in this country by Allen[18] and Joslin[19] with apparent success. By the reduction of carbohydrate, or by the starvation treatment, it is usually possible to make even the severest cases of diabetes aglycosuric, and when this has been attained, then gradually to increase the amount of protein or carbohydrate food until the assimilation limit has been reached.

Saturation Limits.—To avoid error caused by irregular absorption from the intestines, some investigators have recommended the determination of the assimilation limit after intravenous or subcutaneous injections of sugar. But even this refinement in technic has not, as a rule, had the effect of rendering the results of any very evident value as a criterion of the utilization of glucose in the animal body. The reason for this unreliability of the method is mainly that the period of injection of the glucose solution usually occupies only a few minutes, so that it causes

a sudden instead of a very gradual increase in the sugar concentration of the blood, the conditions being quite unlike those which exist during the normal absorption of glucose from the intestine. The mechanism by which the body ordinarily disposes of excessive amounts of glucose absorbed into the portal blood, is not adjusted to operate when the systemic blood is suddenly overcharged with this substance. In the one case the glucose is a foodstuff; in the other, because of its excessive concentration in the blood, it is more or less of a poison. Such results, in other words, merely show us how much glucose can be added at one time to the organism without any overflow into the urine, but they furnish us with no information regarding the power of the organism to utilize a constant though moderate excess of this substance. In the one case it is the "saturation limit," in the other the "utilization limit" of the organism for glucose, that we are really considering.

Consideration of these principles has led Woodyatt, Sansum and Wilder[20] to undertake a thorough reinvestigation of the whole problem of the utilization or, as they prefer to call it, *the tolerance of the body for glucose.* They emphasize the obvious fact that the ability of the organism to utilize glucose "must depend on the rate at which the tissues are able to abstract it from the blood by their combined powers, to burn it, to reduce it into fat or to polymerize it into glycogen." To form any estimate of the combined effect of these processes, we must take into account not only the amount of glucose per unit of body weight (grams per kilogram), but also the rate of injection, for "tolerance must be regarded as a velocity, not as a weight."

Briefly summarized, the conclusions which Woodyatt, etc., have so far drawn from their investigations are as follows: In a normal rabbit, dog, or man, 0.8-0.9 gm. of glucose per kilogram body weight and per hour can be utilized by the organism for an indefinite time without causing glycosuria. When between 0.8 and 2 gm. are injected, a part of the excess appears in the urine, steadily increasing until a maximum is reached, after which the excreted fraction remains constant (at about one-tenth). If more than about 2 grams per kilogram an hour are injected, "a large percentage of all glucose in excess of the 2 gm. per kilogram an hour appears in the urine when constant conditions are once established."

The fact that so much glucose injected intravenously can be used without the appearance of any of it in the urine, indicates a method by which foodstuffs may be supplied to the tissues in cases where, on account of gastrointestinal disturbances, it is impossible to have food absorbed by the usual pathways. The possible value of such a method of treatment in cases of extreme weakness has been tested on laboratory animals by Allen, who states that such injection seems to have a valuable nutri-

tive and strengthening effect. He found, for example, that in cats starved to extreme weakness the injection of a fraction of a gram per kilogram of glucose had an unmistakable strengthening effect, and sometimes appeared to save life. Such results would seem to indicate that in certain cases where blood transfusion is impracticable, glucose infusions should be tried. *Subcutaneous* injection of sugar, either for the purpose of determining the assimilation limit or with the object of supplying foodstuffs parenterally, is impracticable because of the pain and sometimes sloughing produced at the point of injection.

We have devoted no inconsiderable space to a discussion of assimilation limits because of the great interest in diabetic therapy which this procedure has aroused during recent years. We may now turn our attention to a closer analysis of the changes that take place in carbohydrates during their passage through the animal body.

DIGESTION AND ABSORPTION

Digestion.—All digestible carbohydrate taken with the food is converted by the digestive agencies into the monosaccharides, glucose and levulose, as which it is absorbed into the blood of the portal system. To bring about this resolution of carbohydrate into monosaccharides, several enzymes are employed. The first of these is the *ptyalin* of saliva. It is not a very powerful enzyme, being capable of acting only on starches that are in a free state, i. e., not surrounded by a cellulose envelope; but even on free starch, ptyalin displays little of its activity during the time the food is in the mouth. After the food is swallowed and becomes deposited in the fundus of the stomach, there is an interval of time—lasting until hydrochloric acid has been secreted to such an extent as to permit some of the acid to exist in a free state—during which the ptyalin acts on the starch of the swallowed food. During this time the activity of the ptyalin is actually assisted on account of the fact that a slight increase in hydrogen-ion concentration of the digestive mixture accelerates the action of ptyalin.

The product of ptyalin digestion is maltose, a disaccharide composed of two molecules of glucose. On entering the intestine, the carbohydrates therefore exist partly as undigested starch, partly as glucose, and partly as maltose. In the favorable environment of the duodenum a much stronger diastatic enzyme called *amylopsin* very quickly hydrolyzes the starch through dextrine into maltose. The maltose derived from the starch and the unchanged sugars, such as cane sugar, maltose and lactose, which have been taken with the food, unless they are present in very high concentration in the intestinal contents, are not immediately ab-

sorbed into the blood, but become subject to the action of other enzymes contributed by the intestinal juice—namely, the *inverting enzymes,* one of which exists for each of the disaccharides. By their action maltose is converted into two molecules of glucose by the enzyme maltase; lactose, into galactose and glucose by lactase; and cane sugar, into levulose and glucose by invertase. It is interesting to note that in animals whose food does not contain one of those disaccharides, the corresponding inverting enzyme is absent from the intestinal juice. The herbivorous animals, for example, do not take any lactose in their food, and the intestinal juice contains therefore no lactase, although it is present in that of the young animals while still suckling.

A certain amount of carbohydrate becomes attacked by the *intestinal bacteria.* These split the monosaccharides into lower fatty acids and gases, such as methane and carbon dioxide. Besides this obviously destructive process, bacteria also perform a useful function in the digestion of carbohydrates, in that certain strains of them are able to digest cellulose, for which no special enzyme is provided. Bacterial digestion is consequently essential in herbivorous animals; it takes place in the cecum, which is enormously developed for this purpose (page 463).

Absorption.—The glucose and levulose produced by digestion are absorbed into the blood of the portal system. When a very large quantity of a disaccharide, such as cane sugar, is present in the food, a certain amount of the sugar is absorbed unchanged—that is to say, as cane sugar—and appears in the blood, from which, since it is an abnormal constituent, it is excreted unchanged in the urine. This alimentary glycosuria is particularly evident when the sugar is taken without any other food; thus, after taking cane sugar in an amount corresponding to 5 grams per kilogram body weight, it was found in one and a half hours afterward that the urine of ten out of seventeen healthy individuals contained cane sugar. The urine of three of these men, however, also contained invert sugar—that is, dextrose and levulose. Cane sugar continued to be excreted for from six to seven hours.

The Sugar Level in the Blood.—While no absorption of sugar is going on, the percentage of this substance in the blood of the portal vein is the same as that in the systemic circulation. During absorption the former becomes perceptibly raised—to what extent we can not say—and in the latter a less marked increase of sugar concentration is usually detectable. Evidently, then, between the point at which the sugar is absorbed and the blood of the systemic circulation, some barrier exists which holds back some of the excess of absorbed sugar. We have very inaccurate information as to how efficiently these barriers hold back the excess of absorbed glucose because of the technical difficulty in collecting blood

from the portal vein without serious disturbance to the animal. Indeed, the only way by which the problem has been accurately studied is by comparing the blood of the portal circulation with that of the systemic circulation during the injection of a solution of dextrose into one of the smaller branches of the portal vein.[21] In such experiments it has been found that the percentage of sugar is a little less in the blood of the abdominal vena cava than in that of the portal vein, and is still less in the blood of the systemic veins, such as the femoral—results which justify the conclusion that the barriers responsible for taking out some of the absorbed sugar from the blood exist in the liver and in the muscles. The curve in Fig. 189 will illustrate to what extent the mechanism operates.

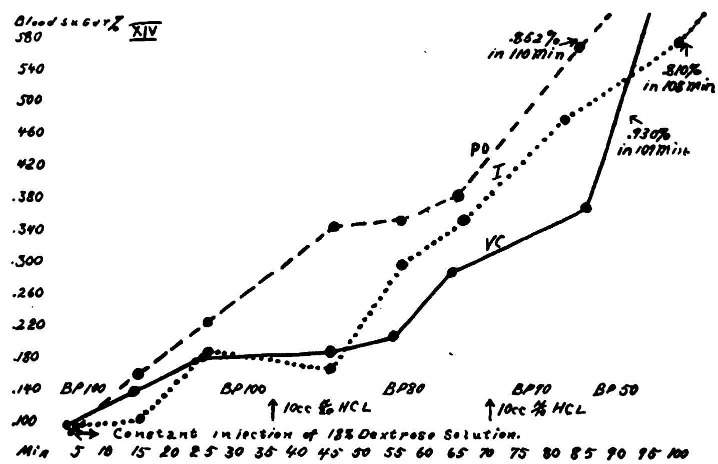


Fig. 189.—Curves showing the percentage of glucose in blood after a constant injection of an 18 per cent solution into a mesenteric vein. *V.C.*, vena cava, continuous line; *P.D.*, pancreaticoduodenal vein, broken line; *I*, iliac, dotted line.

It will be observed that, so far as can be judged from changes in the concentration of sugar in the blood, the sugar-retaining power of the liver is about equal to that of the muscles—a conclusion which is, however, contrary to the usually accepted one that the liver has such pronounced sugar-retaining powers that under ordinary circumstances it removes from the portal blood all the excess of sugar added to it by absorption and which is not required by the organism.

One objection which may properly be made to these observations is that the animals on which they were made were under anesthesia, and that the anesthetic had a paralyzing effect on the sugar-retaining power of the liver. In view of this criticism it is important to examine the results obtained on animals that are not under the influence of anes-

thesia. Such observations have been made on rabbits, and a few on man himself. By collecting blood from the ear veins of rabbits, it has been found that, after giving from two to ten grams of glucose by stomach, the glucose concentration of the systemic blood begins to rise in fifteen minutes, attaining a maximum in about an hour and then returning to the normal level in about three hours.

Similar results have been obtained by examination of the venous blood in man. After giving 100 grams of glucose by mouth, for example, there is commonly an increase in blood sugar amounting to from 30 to 34 per cent of the normal and lasting for from one to four hours. The existence of this *postprandial hyperglycemia,* as we may call it, indicates that the sugar-retaining powers of the liver and muscles are not sufficiently developed to prevent the accumulation of some of the absorbed sugar in the systemic blood. Whenever this increase exceeds a certain limit, some of the sugar begins to escape through the kidney into the urine, producing glycosuria—*postprandial glycosuria.* The percentage of blood sugar above which glycosuria occurs is, in the case of man, probably about 0.10 to 0.11 gm. per cent. After damage to the kidney, as in nephritis, or in long-standing cases of mild diabetes, the percentage may probably rise considerably higher in the blood without evidence of glycosuria.

Value of Blood Examination in Diagnosis of Diabetes.—The determination of the amount of ingested carbohydrate required to bring about postprandial glycosuria constitutes, as we have already seen, the so-called assimilation limit for sugar, which is often taken as an index of the sugar-metabolizing power of the organism. It is evident, however, that the time of onset, and the extent and duration of postprandial hyperglycemia must serve as a more certain index of the sugar-retaining power of the liver and muscles; and now that a simple and rapid clinical method exists (Lewis-Benedict method) for the accurate determination of sugar in small quantities of blood, there is no reason why this index should not be used for the detection of failing powers to metabolize carbohydrate.

In no disease, probably not even in tuberculosis, is it more important than in diabetes that an early diagnosis should be made. Thus, if we find that the postprandial hyperglycemia after a certain amount of carbohydrate develops to an unusually high degree and persists for an unusual length of time, we are justified in curtailing the carbohydrate supply so as to hold these values down to their level in normal individuals. It is almost certain that the first sign of diabetes is an unusual degree and duration of postprandial hyperglycemia. At first the excess of sugar leads to no damage and it is insufficient to cause any evident glycosuria, although it is quite likely that if the urine in such individuals were collected at very frequent intervals after eating carbohydrate-rich food, glucose would be

found present in at least some of the specimens. In incipient diabetes, however, the condition progresses, until the postprandial hyperglycemia after one meal has not become entirely replaced before the next is taken, so that the increase in sugar produced by the second meal becomes superadded on that following the first meal. The curve of blood sugar rises ever higher and higher, until at last permanent hyperglycemia is established, or rather the normal level from which the postprandial rise occurs has become permanently raised, so that in blood collected at *any* time a higher percentage of sugar is found.

The Relationship Between the Sugar Concentration of the Blood and the Occurrence of Glycosuria.—Claude Bernard first pointed out that the percentage of sugar in the blood may rise considerably above its normal level without the appearance of any of the sugar in the urine, or at least without a sufficient amount to give the usual tests for sugar. Even when this limit is reached, as we have seen, the sugar which appears is not all of the excess but only a small part of it. This overflow hypothesis, as it is called, has not been universally accepted because of the many results which are not in conformity with it. Many of these exceptional results have been explained as due to alterations in the permeability of the kidney for sugar, and in general it is probably safe to accept Claude Bernard's hypothesis with certain reservations.

Strong support has been lent to a modified form of the hypothesis by the recent work of Woodyatt and his collaborators, who have shown by continuous intravenous glucose injections that as much as 0.8 gm. of glucose per kilo body weight can be injected during an hour into an animal without any glycosuria, although under such conditions a very distinct increase occurs in the percentage of sugar in the blood.

To explain the failure of glucose to pass into the urine under normal conditions, it has been supposed by several investigators that the glucose exists in some form of chemical combination in the blood. This compound is believed to behave like a colloid. One of the recent supporters of this view is Allen, who has observed that, when glucose is injected intravenously, it causes diuresis as well as glycosuria; whereas glucose injected subcutaneously or taken by mouth causes neither of these conditions to become developed; indeed it causes for some time after the administration of the sugar a distinct anuria. To explain these differences in behavior between glucose administered intravenously and that taken in other ways, it is supposed that the glucose molecule in passing through the intervening wall of the capillaries combines with some substance to form a compound which becomes available for incorporation into and utilization by the tissues, glucose in a free state being incapable of utilization. This compound is supposed to be of a colloidal nature, and the substance which

combines with glucose to form it is believed to be related to the internal secretion of the pancreas (see page 676).

The difficulty in explaining why the glucose of the blood does not constantly leak into the kidney is, however, the only evidence upon which the hypothesis of a blood sugar compound rests. No chemical evidence can be offered in support of such a view. On the contrary, all experimental work indicates that the sugar exists in a free state; but unfortunately even this evidence is not convincing. Thus, it has been found that, when specimens of perfectly fresh blood are placed in a series of dialyzer sacs suspended in isotonic saline solutions, each solution containing a slightly different percentage of glucose, diffusion of glucose, in one or other direction, occurs in all of them save one—namely, that in which the percentage of glucose in the fluid outside the dialyzer is exactly equal to the total sugar content of the blood. Such a result can be explained only by assuming that all of the sugar in the blood exists in a freely diffusible state. In its general nature this experiment is analogous to that by which the tension or partial pressure of CO_2 is determined in blood (see page 338).

It has been assumed by many clinicians that glycosuria may sometimes become developed because the kidney fails to hold back the blood sugar even when the percentage is not above the normal—so-called *renal diabetes.* For the diagnosis of this condition a comparison must be made between the sugar concentration of the blood and that of the urine. In order to do this at least two samples of blood must be taken, one of them at the beginning and the other at the end of a period during which urine is being collected. Merely to find that one sample of blood collected before or after or during the period of urine collection contains a normal percentage of sugar, does not necessarily indicate that at some other period while the urine was being produced a temporary hyperglycemia may not have existed.

CHAPTER LXXV

THE METABOLISM OF THE CARBOHYDRATES (Cont'd)

FATE OF ABSORBED GLUCOSE. GLUCONEOGENESIS

We may now consider what becomes of the sugar that is retained by the liver and muscles. Two things may happen to it: It may become stored, or it may become oxidized or split up. Of these processes, storage occurs in both the liver and muscles, whereas oxidation occurs mainly if not entirely in the muscles, although a certain amount of splitting of the glucose molecule may also occur in the liver.

Storage of Sugar.—For the present we shall consider the process of storage of sugar and defer a consideration of its utilization until after we have studied, not only the nature of the process by which the storage occurs, but also the immediate destiny of the stored sugar. The storage of sugar by the liver is brought about by its conversion into a polysaccharide called *glycogen.* After an animal has been absorbing large quantities of glucose, an acidified watery extract of a portion of liver made immediately after death will be found to contain no more sugar than that of a normal liver. On the other hand, it will be observed that the extract is highly opalescent and yields on the addition of alcohol a copious precipitate, which on further purification can readily be shown to consist of a polysaccharide—that is to say, of a starch-like substance which on hydrolysis with mineral acid becomes entirely converted into sugar. If instead of removing the liver immediately after death, it is allowed to stand for some time, the yield of glycogen will greatly diminish, and in its place will appear large quantities of glucose, indicating that some enzyme must exist which attacks the glycogen after death and converts it into sugar This enzyme is called *glycogenase.* The existence of *postmortem glycogenolysis,* as it is called, would seem to indicate that during life also there is a constant tendency for the glycogen in the liver to be attacked by glycogenase, but that this is prevented by conditions which depend on the vital integrity of the liver cell. It is evident that if anything should happen during life to interfere with this inhibiting influence, the glycogen will become converted into glucose, which on escaping into the blood will produce hyperglycemia and glycosuria.

Sources of Glycogen.—In studying the source of sugar in the animal body it is important therefore that we should first of all know exactly the

conditions under which glycogen may be formed in the liver; that is, whether it is formed from absorbed sugar alone, or from other substances also, such as protein and fat. The importance of such knowledge rests in the fact that in severe diabetes, sugar still continues to be added to the blood, although no sugar is being taken with the food. To check the hyperglycemia in such cases it becomes necessary, therefore, to curtail the diet not only with regard to its carbohydrate content, but also with regard to whatever other foodstuff may be capable of causing glycogen formation. The practical question therefore is, What are these foodstuffs?

There are two methods by which the problem may be investigated. The first, which we may call the direct method, consists in rendering the liver free of glycogen and then some time afterward feeding the animal with the foodstuff in question, afterward killing it and examining the liver for glycogen. The other, which we may call the indirect method, consists in first of all rendering the animal incapable of oxidizing glucose—that is, making it diabetic—and then proceeding to see whether the ingestion of a given foodstuff causes an increase in the sugar excretion in the urine. The methods for rendering an animal experimentally diabetic will be considered later; for the present it is important to note that, if a diabetic animal excretes more glucose while fed on a given foodstuff, we may infer that the normal animal would convert it into glycogen.

The results of the direct method are much less reliable than those of the indirect for the reason that it is extremely difficult to remove all traces of glycogen from the liver. The methods employed for this purpose have consisted in: (1) starvation of the animal; (2) muscular exercise; (3) exercise and starvation combined; and (4) the production of certain forms of experimental diabetes—for example, that produced by phlorhizin. Starvation alone is unsatisfactory, for it has been found that, although at certain stages of this condition the liver may become almost entirely free from any trace of glycogen, at a later stage glycogen may again make its appearance. It is therefore most difficult to decide at what stage in starvation the animal should be considered as glycogen-free.

If the starving animal is made to perform muscular exercise, complete removal of glycogen from the liver can be depended upon. The exercise may be produced by the administration of strychnine in such dosage as just to produce convulsions of the voluntary muscles without permanent contraction of those of respiration. The most useful method, however, consists in starving the animal for a few days and then placing it in a cold, damp room, after giving it a cold bath. The evaporation of moisture from the surface so cools the body down that the glycogen store all becomes used up in the attempt to supply fuel for the production of

sufficient heat to maintain the body temperature. This method can be rendered still more certain in effecting a removal of all carbohydrate from the body by giving the animal phlorhizin every eight hours. This drug, as we shall see, renders the animal diabetic.

After removing the glycogen, further deposition in the liver can be readily shown to occur when any of the ordinary sugars or starches are given as food. It does not occur, however, when chemical substances closely related to ordinary sugar, such as the wood sugars (pentoses) or the alcohols and acids corresponding to dextrose, are contained in the diet. Nor does it occur with cellulose or with inulin, a polysaccharide built up from pentose sugar. When proteins are fed the results are not so definite, although many observers have claimed that glycogen is formed. With fat, on the other hand, no glycogen formation can be shown to occur, although we know that a trace of carbohydrate must be formed out of the glycerine of the fat molecule.

The results of the direct method, even when the conditions are perfectly controlled, are very unreliable, because any new sugar produced by the ingested substance instead of being stored as glycogen may be directly used by the tissues as it is formed. Where only a slight degree of gluconeogenesis is occurring, it is not likely that any of the glucose will be retained in the body as glycogen.

The methods employed for producing experimental diabetes in investigation of these problems by the **indirect method** are (1) the entire removal of the pancreas, and (2) the continuous administration of the drug phlorhizin. The animal rendered diabetic by either of these methods is first of all observed for several days to determine the normal daily excretion of sugar. At the same time the nitrogen excretion for the day is determined, the ratio between the total nitrogen and the glucose being known as G to N ratio, and being about 1 to 3.65 when complete diabetes has become established. The foodstuff in question is then fed to the animal, and the amount of extra glucose excreted thereby is taken to represent that which has been derived from the ingested food. By this method it has been possible to show that, not only the above mentioned carbohydrates, but protein as well produce a very considerable quantity of glucose in the animal body. Fats, however, yield only negative results.

The indirect method has another great advantage over the direct in that the results are much more *quantitative* in character; for example, Lusk and his pupils have been able to determine the amount of glucose which can be produced by feeding certain of the building stones of the protein molecule. The great practical importance of such results in

the therapy of diabetes makes it advisable for us to go into the subject a little more in detail here.

Dogs are rendered diabetic by phlorhizin followed by a cold bath and exposure in a cold room. When all of the original glycogen in the body has been got rid of, as evidenced by the constancy of the G to N ratio in the daily quantities of urine excreted, the substance under investigation is fed. If this substance contains no nitrogen and causes no change in the nitrogen excretion, any increase in that of glucose must obviously represent the extent to which the substance has become converted into this sugar. On the other hand, if the substance itself contains nitrogen, or if it causes a change in the excretion of nitrogen, it becomes necessary to calculate how much of the excreted glucose might have been derived from the body protein, assuming that this can form glucose, and how much from the administered substance.*

From the results of this method it has been an easy matter to show that the following substances are converted in the animal body into glucose: (1) *Glycol aldehyde* ($CH_2OH - CHO$). By placing three molecules of this substance together, a hexose molecule will be produced, a synthesis which can be accomplished in the chemical laboratory. Glycol aldehyde may be formed in normal metabolism out of glycocoll (CH_2NH_2COOH).

(2) *Glycerol* ($CH_2OH - CHOH - CH_2OH$) may also readily be converted into hexose in the laboratory, the possible intermediary products being dioxyacetone ($CH_2OH - CO - CH_2OH$) and glyceric aldehyde ($CH_2OH - CHOH - CHO$). Two molecules of either of these may be polymerized to form a hexose molecule, and when this process occurs in the animal body, the hexose formed is glucose.

(3) *Lactic acid* ($CH_3CHOH - COOH$) is completely converted to dextrose in the diabetic animal, and the process must involve both a rearrangement of the molecule and subsequent polymerization. The related substance, propyl alcohol ($CH_3 - CH_2 - CH_2OH$) is also converted into glucose in the phlorhizinized dog. As to the exact nature of the chemical changes which occur as intermediary steps in the conversion of these substances into glucose, we are not as yet certain, but a step has been made in the discovery that a substance called methylglyoxal (CH_3COCHO) may be obtained from lactic acid and also from glucose, and that this substance is converted into glucose when it is administered to phlorhizinized dogs. We shall find later an important role for this substance

*This calculation is made as follows: The amount of nitrogen in the administered substance is deducted from the nitrogen excretion, and the difference, which must represent the nitrogen of the body protein, is multiplied by the G to N ratio which prevailed on the day previous to that on which the substance was fed. We obtain in this way the glucose derived from the body. The glucose coming from the administered substance can then be ascertained by deducting that derived from the body protein from the total glucose excretion.

in the case of fat metabolism, and it can also readily be produced during the intermediary breakdown of certain of the protein building-stones, such for example as alanine (CH_3CHNH_2COOH).

These chemical possibilities regarding the nature of the substances that serve as stepping stones between the above sugar-forming substances and sugar itself may be translated later into probabilities on account of the discovery that the enzymes exist in various tissues which are involved in converting methylglyoxal into lactic acid:

$$\begin{array}{ccc} CH_3 & & CH_3 \\ | & & | \\ CO & + \quad H_2 \rightarrow & HCOH \\ | & O \leftarrow & | \\ CHO & & COOH \\ \text{(methylglyoxal)} & & \text{(lactic acid)} \end{array}$$

These enzymes are called *glyoxalases,* and since the reactions which they mediate are undoubtedly reversible in character, it is probable that the conversion into sugar of lactic acid and alanine—to take those two as among the commonest of the sugar precursors of the animal body—occurs according to the following equation:

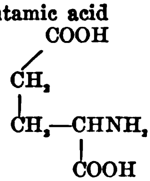
$$\begin{array}{lll} CH_3CHNH_2COOH \searrow & & \\ \quad \text{(alanine)} & CH_3COCHO \rightarrow C_6H_{12}O_6 & \\ CH_3CHOHCOOH \nearrow & & \\ \quad \text{(lactic acid)} & \text{(methylglyoxal)} \quad \text{(hexose)} & \end{array}$$

The unique position of methylglyoxal, besides explaining the known resolutions of protein and fat and carbohydrate in intermediary metabolism, is also of importance in explaining the synthetic production of glucose from fructose (or levulose). Fructose will first of all become converted into methylglyoxal radicles, and these will then become synthesized into glucose.

The theory of the conversion of glucose into lactic acid as a stepping stone in the metabolism of carbohydrate meets with one objection; namely, that the lactic acid is not produced from carbohydrate in the organism, except in cases where there is oxygen deficiency or excess of alkali in the tissue fluids.

Coming now to the amino acids, which, it will be remembered represent the building stones of the protein molecule, it has been found that glycocoll, alanine, and aspartic and glutamic acids increase the glucose excretion when given to phlorhizinized dogs, whereas leucine and tyrosine have no such action. By the method described above, it is possible to determine the exact proportion of the carbon of each of those amino acids which becomes converted to glucose. This is shown in the accompanying table.

TWENTY GRAMS OF THE VARIOUS AMINO BODIES WERE GIVEN TO PHLORHIZIN-DIABETIC DOGS

ACID AND FORMULA	AVERAGE AMOUNT OF GLUCOSE PRODUCED IN BODY	PROBABLE CHANGE	GLUCOSE THAT WOULD BE PRODUCED BY CHANGE
Glycocoll CH_2NH_2COOH	13.43 (five dogs, one gave 15.77)	All C converted to glucose	16.00
i. alanine CH_3CHNH_2COOH	18.77 (two dogs)	"	20.22
Aspartic acid $COOH—CH_2—CHNH_2—COOH$	12.42 (four dogs)	Three of the four C atoms converted to glucose	13.52
Glutamic acid 	13.31	Three of the five C atoms converted to glucose	12.24

It is of further interest to point out that these four amino acids constitute about 26 per cent of all the amino acids in flesh protein, and that the total yield of glucose from them could be 26.3 grams; thus accounting for nearly one half of the 66 grams which a diabetic animal produces from 100 grams of flesh.

Gluconeogenesis in Normal Animals.—Although it has been clearly shown by the indirect method that not only protein but its decomposition products as well, can be readily converted into glucose, yet this does not necessarily indicate that a similar conversion occurs in the nondiabetic animal. That such is the case, however, can be shown in various ways. Thus, at the end of a period of long starvation considerable quantities of glycogen are quite commonly found in the body, and the blood sugar, although lower than normal, never entirely disappears. Now, since no carbohydrate is being ingested, and the body stores of this foodstuff become exhausted early during starvation (cf. page 663), it is evident that the carbohydrate must be produced from the protein of the animal's body. A still more convincing experiment can be conducted by producing strychnine convulsions in a starving animal. If the animal is killed after the convulsions have lasted for a certain time, the tissues will be found almost if not entirely free of glycogen, but if the convulsions are made to disappear by giving chloral and the animal allowed to sleep for some time before killing it, glycogen again accumulates in the body. This glycogen must have been manufactured out of noncarbohydrate material.

Corroborative evidence of a somewhat different nature is furnished by

an examination of the respiratory quotient, which, it will be remembered (page 547), varies according to the nature of the foodstuff or body constituent that is undergoing metabolism at the time, being about 1 with carbohydrate and about 0.7 with protein. If the quotient is observed during starvation, it will often be found to fall below 0.7, a figure which can be explained only by assuming that oxygen has been retained in the body beyond the quantity which is necessary for immediate purposes of oxidation (cf. equations on page 548).

Since it is known that this retained oxygen can not exist in the body in a free state it must be concluded that it has become incorporated into substances having a high oxygen content. Such would be the case if protein or fat, which contains only from 12 to 20 per cent of oxygen, were converted to carbohydrate, which contains about 53 per cent. Utilization of inhaled oxygen for this purpose, as we have seen, becomes very striking in the case of hibernating animals during the winter sleep.

CHAPTER LXXVI

THE METABOLISM OF THE CARBOHYDRATES (Cont'd)

FATE OF GLYCOGEN

Having become familiar with the sources from which glycogen may be derived, we may now proceed to study the fate of the glycogen found in the liver cells and in the muscles. For the present we shall confine our attention to the glycogen of the liver. If a portion of liver removed from a well-fed animal is examined microscopically after staining either with iodine or with carmine by Best's method, it will be found that the cells of the lobules are filled with glycogen except for the nuclei, which are free from this substance. If, on the other hand, the liver is from an animal that has not been recently fed, the lobules will contain no glycogen except for an area bordering on the central vein and perhaps a narrow strip at the periphery of the lobule. When it is present the relative amount of glycogen in different lobules, as determined chemically, is the same over the entire liver—that is to say, no one lobe is richer in this substance than another. Nothing definite is known as to how the glycogen is held in the protoplasm of the cells, although some histologists suggest that it is combined with a sustentacular material especially provided for this purpose.

The glycogen stored in the liver is gradually given up to the blood of the hepatic vein at such a rate as to maintain in the blood of the systemic circulation a more or less constant percentage of glucose. Under ordinary conditions this process of *glycogenolysis* is relatively slow, but when the requirements of the organism for fuel become increased, as during muscular exercise, it becomes very rapid. The glycogenic function of the liver appears therefore to exist, in part at least, for the purpose of preventing the flooding of the blood of the systemic circulation with excess of sugar during absorption from the intestine and of maintaining the normal percentage at other times. This function is analogous to that occurring in plants, in which the sugar produced in the leaves, if not immediately required, is transported to various parts of the plant and there converted into starch, which, when the plant requires it, as during new growth, may again become transformed into glucose.

The agency converting the glycogen into glucose is the diastatic

enzyme *glycogenase,* which is present, not only in the liver cell, but also in the blood and lymph. It is a difficult matter to explain why glycogen should be able to exist at all in the liver cells in the presence of this powerful enzyme. The following possibilities may be considered: (1) That glycogenase does not really exist in the living liver cells, but is a postmortem product; (2) that, although present, glycogenase is prevented from acting on the glycogen in the living liver cell on account of the latter being protected from its influence by combination with the sustentacular substance; or (3) that some chemical substance in the liver cell prevents the glycogenase from acting on the glycogen—an anti-glycogenase. Since the removal of any one of these inhibiting influences would cause glycogenolysis to become excessive, and so bring about hyperglycemia, it is important, in searching for the possible causes of this condition, to examine the evidence that has been brought forward in support of each of these views.

Against the view that glycogenase is a postmortem product may be cited the very rapid conversion into glucose that occurs when glycogen is added to living blood, as by injecting some into a vein. On account of the active glycogenolytic action of blood, it has been suggested that during life glycogen does not become transformed into glucose until after it has been discharged into the blood from the liver cell. When increased sugar must be mobilized, glycogen passes unchanged, or perhaps as some dextrine, into the blood and lymph of the liver capillaries and lymphatics, the glycogenase of which converts it into glucose, the conversion being so rapid that, by the time the blood has traveled from the liver through the heart and pulmonary vessels to the arteries, all the glycogen has already become transformed into glucose. Postmortem glycogenolysis, according to this view, is due to the opposite occurrence—the transference of glycogenase from the blood into the liver cell. Some facts supporting this view are as follows: (1) It has been found that the amount of free glucose in the blood of the vena cava is sometimes less than in that collected simultaneously from the carotid artery. (2) After giving certain substances, such as phosphorus or peptone, there is distinct diminution in the amount of glycogen in the liver, accompanied, however, by no increase in the amount of glucose in the blood. And (3) if the liver of an animal that has been rendered diabetic by stimulation of the splanchnic nerve or by puncture of the floor of the fourth ventricle is examined microscopically, after staining by the carmine method, masses of stained glycogen can be found present in the capillaries (sinusoids) that lie among the liver cells.

According to the second view, the glycogen is removed from the influence of the intrahepatic glycogenase on account of its combination

with a sustentacular material. By disrupting this combination and thus exposing the glycogen to the action of glycogenase, glycogenolysis will occur. We may call this the mechanical hypothesis and it deserves serious consideration, for it has been shown that very little postmortem glycogenolysis occurs in the intact liver of frogs in winter,—even though at this time the organ contains an excess of glycogen,—but becomes marked when the liver is broken down by mechanical means.

The third view depends on the well-known fact that enzyme activities become most markedly altered by slight changes in the chemical nature of the environment in which they act. Diastatic enzymes are particularly susceptible to the reaction (C_H) of their environment, a very slight degree of acidity favoring and a trace of alkalinity markedly depressing their activities. That a tendency to increasing acidity in the liver cells may retard the formation of glycogen is suggested by the depressing effect produced on the assimilation limit of sugars by administering acids, and by the observation that postmortem glycogenolysis becomes marked in proportion as the dying liver becomes acid in reaction. It might be thought then that glycogenolysis in the liver cell could be set up by the local production of a certain amount of acid. Such a liberation of free acid could be brought about by a curtailment in the arterial blood supply of the hepatic cell, producing a local accumulation either of carbonic or of other less completely oxidized acids (e. g., lactic). It may be that asphyxia causes hyperglycemia by such a mechanism. Vasoconstriction and consequent curtailment of arterial blood supply occurs in the liver when the hepatic nerves are stimulated, and it is possible that the glycogenolysis which is also set up by such stimulation is due to the appearance of acids. The accelerating effect of epinephrine on glycogenolysis might also be explained as due to limitation of blood supply on account of vasoconstriction.

THE REGULATION OF THE BLOOD SUGAR LEVEL

The level at which the concentration of sugar in the systemic blood is maintained represents the balance between two opposing factors: (1) the consumption of glucose by the tissues, and (2) the production of glucose by the liver. Since this is the most readily oxidizable of all the proximate principles of food (page 652), muscular activity causes large quantities of it to be consumed, so that its concentration in the blood tends to fall below the physiologic level, a tendency which is immediately met by an increased discharge of glucose from the liver. The question therefore arises as to *how the muscles or other tissues transmit their requirements for glucose to the liver.* There are two

possible ways by which this could be done: (1) by means of a nervous reflex, or (2) by changes in the composition of the blood, either with regard to the percentage of sugar itself or because of the appearance in it of decomposition products of glucose or of some special exciting agent or hormone.

In order to ascertain the relative importance of these methods of correlation between the places of supply and demand of glucose in the normal animal, it is necessary to investigate the conditions under which an excessive discharge of glucose occurs as a result of overstimulation of the nervous control, or because of the presence of exciting substances (hormones) in the blood. The glycogenic function can be excited through the nervous system in a variety of ways so as to produce hyperglycemia and glycosuria. This constitutes one form of *experimental diabetes.* In laboratory animals mechanical irritation of the medulla oblongata and stimulation of the great splanchnic nerves act in this way. Similar stimulation may also occur under certain conditions in man. Excitation as a result of changes in the composition of the blood can be produced experimentally by certain drugs (phlorhizin), or by the removal of certain of the ductless glands or the injection of extracts prepared from them, such as epinephrine.

Nerve Control and the Nervous Forms of Experimental Diabetes.—The simplest experimental condition which illustrates the relationship between the nervous system and the blood sugar is *electrical stimulation of the great splanchnic nerve* in animals in which, by previous feeding with carbohydrates, a large amount of glycogen has been deposited in the liver. By examination of the blood as it is discharged into the vena cava from the hepatic veins, the increase in blood sugar is very evident in from five to ten minutes after the first application of the stimulus; but it is not until later that a general hyperglycemia becomes established. The conclusion which we may draw from these results is that the splanchnic nerve contains efferent fibers controlling the rate at which glycogen becomes converted to glucose in the liver. The center from which these fibers originate is situated somewhere in the medulla oblongata, for the irritation that is set up by puncturing this portion of the nervous system with a needle yields results similar to those which follow splanchnic stimulation. This *"glycogenic" or diabetic center,* as it has been called, must be provided with afferent impulses. Such impulses have indeed been described in the vagus nerves, but their demonstration is by no means an easy matter on account of the disturbance in the respiratory movements coincidently produced by the stimulation. The changes that such disturbances bring about in the aeration of the

blood may in themselves be responsible for the hyperglycemia (see page 332). It can at least be said that when the respiratory disturbances are guarded against, as by intratracheal insufflation of oxygen, vagal hyperglycemia is much less marked, if not entirely absent. But this question awaits more thorough investigation.

The increased glycogenolysis which results from stimulation of the efferent fibers in the splanchnic nerves may depend either on a direct control exercised over the glycogenic functions of the hepatic cells, or on the discharge into the blood of some hormone which excites the glycogenolytic process. It must furthermore not be lost sight of that the glycogenolysis may be secondary to local asphyxial conditions in the liver cells resulting from vasoconstriction. From their anatomic position, the adrenals are to be thought of as the source of the hormone, and evidence that splanchnic hyperglycemia is due to hypersecretion from these glands has seemed to be furnished by the fact that after they are extirpated splanchnic stimulation no longer produces hyperglycemia, neither, indeed, does puncture of the medulla. There is also no doubt that the nervous system, acting by way of the splanchnic nerves, does exercise a control over the discharge of the internal secretion of the adrenal glands and that extracts of the gland, which we must suppose act in the same way as the internal secretion, cause hyperglycemia when injected intravenously (epinephine hyperglycemia and glycosuria).

But on theoretical grounds alone, certain difficulties immediately present themselves in accepting this as the mechanism by which the nervous system controls the sugar output of the liver, for if increased sugar formation in the liver is dependent on a discharge of epinephrine, the question may be asked why this secretion should be caused to traverse the entire circulation before reaching the liver.

There are, besides, certain experimental facts which do not conform with such a view. Thus, after complete severance of the hepatic plexus of nerves, stimulation of the splanchnic nerve does not cause the usual degree of hyperglycemia, whereas electric stimulation of the peripheral end of the cut plexus does cause it. On the one hand, therefore, there is evidence that stimulation of the efferent nerve path above the level of the adrenals has no effect on the sugar production of the liver in the absence of these glands; and on the other, we see that when they are present, stimulation of the nerve supply of the liver is effective, even though the point of stimulation is beyond them. There is but one conclusion that we may draw—namely, that the functional integrity of the efferent nerve-fibers that control the glycogenolytic process of the liver depends on the presence of the adrenals, very probably because of the hormone which the glands secrete into the blood. This conclusion is

corroborated by the fact that stimulation of the hepatic plexus, even with a strong electric current, some time after complete removal of both adrenals, is not followed by the usual degree of excitement of the glycogenolytic process.

These experiments demonstrate an important relationship between the nervous control, and at least one form of hormone control, of the sugar output of the liver. They indicate that when a sudden increase of blood sugar is required, the glycogenic center sends out impulses which not only directly excite the breakdown of glycogen in the hepatic cells, but also simultaneously influence the adrenals in such a manner as to produce more epinephrine in the blood and so augment the action of the nerve impulse.

We are as yet quite in the dark as to the mechanism by which the nerve impulses or the hormone brings about increased glycogenolysis. It must consist of a removal of the influence that prevents glycogenolysis from occurring in the normal liver, for it has been shown by direct observation that there is no increase in the amount of glycogenase present in extracts of the liver removed from diabetic animals over that present in extracts of the liver of normal animals. The possible nature of this influence has already been discussed (page 669). The change may consist either in a loosening of the combination between the glycogen and the protoplasm of the liver cell, or in a removal of the chemical influence that ordinarily prevents the glycogenase from attacking the glycogen. In the former case the glycogen liberated from its union with the sustentacular substances would either become attacked by the glycogenase present in the liver cell itself or it would first of all migrate, as glycogen, into the blood capillaries and there be attacked by the blood glycogenase. Evidence for the possibility of the occurrence of such a process has already been given (page 670). The chemical change referred to under the second possibility might consist in an alteration in the hydrogen-ion concentration of the liver cell, a change, however, which for obvious reasons it is impossible to investigate.

Nervous Diabetes in Man.—The main interest attaching to the investigation of these nervous forms of experimental diabetes depends on the insight which they afford us into the nature of the mechanism by which a prompt mobilization of glucose may be brought about in the normal animal. There is also some evidence that a relationship may exist between certain of the clinical varieties of the disease in man and repeated excitation of glycogenolysis brought about by nerve stimulation. Increased glucose output from the liver as a result of nerve excitation may be a normal process, but there is reason to believe that frequent repetition of this process tends to induce a permanent rise in the glucose

level of the blood and therefore a tendency to diabetes. There have recently been collected several facts which lend some support to this view. The frequent occurrence of diabetes in those predisposed by inheritance to neurotic conditions, or in those whose daily habits entail much nerve strain, and the aggravation of the symptoms which is likely to follow when a diabetic patient experiences some nervous shock, all point in this direction.

Diabetes is common in locomotive engineers and in the captains of ocean liners—that is, in men who in the performance of their daily duties are frequently put under a severe nerve strain. It is apparently increasing in men engaged in occupations that demand mental concentration and strain, such as in professional and business work. Cannon[23] found glycosuria in four out of nine students after a severe examination, but only in one of them after an easier examination.* In the urine of twenty-four members of a famous football squad, sugar was found present in twelve immediately after a keenly contested game. Anxiety and excitement were responsible for its appearance, for five of the twelve were substitutes who did not get into the game.

Although these nervous conditions, by excitement of hepatic glycogenolysis, produce at first nothing more than an excessive discharge of sugar into the blood—a condition which is exactly duplicated in our laboratory experiments by stimulation of the nerve supply of the liver—their repetition may gradually lead to the development of a permanent form of hyperglycemia. To prevent the repetition of these transient hyperglycemias must be one of our aims in the treatment of early stages of the disease.

Although there can be no doubt that the glycogenic function of the liver is subject to nerve control, it is probable that its control by hormones is of equal if not greater importance. This dual control of a glandular mechanism is by no means unique for the glycogenic function, for we have already seen it to exist in the case of the gastric glands and the pancreas, and it is probable that it also exists in the case of the thyroid. It may well be that the nerve control of the glycogenic function has to do only with those transitory changes in sugar production that would be demanded by sudden activities of muscle, and that the hormone control has to do with the more permanent process of building up and breaking down of glycogen to meet the general metabolic requirements of the tissues.

*We have been unable to confirm this observation even though the examinations were made unusually "nerve-racking."

HORMONE CONTROL AND PERMANENT DIABETES

Nervous excitation can explain only transitory increases in blood sugar, the more permanent hyperglycemias being dependent upon some disturbance in the hormone control of carbohydrate utilization. This disturbance is a much more serious affair than that produced by nervous excitation. In the latter case the hyperglycemia ceases whenever all of the glycogen stores of the liver have been exhausted; whereas a disturbance in the hormone control, besides causing as its first step a breakdown of all the available glycogen, goes on to cause a production of sugar out of protein. A process of gluconeogenesis (new formation of glucose) becomes superadded on one of glycogenolysis.

To ascertain *the nature of this hormone* and the mechanism of its action has been the object of most of the researches on those forms of diabetes produced by changes in certain of the ductless glands. The following possibilities may be considered: (1) that it is the concentration of glucose in the blood; (2) that it is the presence in the blood of decomposition products of glucose; (3) that it is due to a hormone produced from some ductless gland. Regarding the first of these possibilities, it is supposed that the mechanism involved in the adjustment between the blood sugar and the glycogenic function is one explicable on the basis of the law of mass action; namely, that glycogen becomes converted into glucose whenever the blood flowing to the liver contains less than its normal concentration of glucose, and conversely, when this blood contains an excess of glucose, as during absorption, a glycogen-building process takes place. Although there can be little doubt that the process of glycogen formation or destruction will depend to a certain extent upon the amount of glucose present in the blood flowing to the liver cells, yet it is impossible that this can be an important means in the control that exists between sugar production by the liver and sugar consumption by the tissues, because the sugar that is added to the portal blood during absorption would mask any depletion caused by sugar consumption in the tissues.

The second possibility—that the hormone is some decomposition product of glucose—would appear to have some support, if we consider this hormone to be an acid product (carbon dioxide or lactic) produced by sugar metabolism, for it is known that an increase in the hydrogen-ion concentration of the blood flowing to the liver cells excites a glycogenolysis. As we have already seen, however, it is difficult to secure experimental evidence, in anesthetized animals at least, that glycogenolytic activity is readily excited in this way.

The third possibility—that some specific hormone may exist in the

blood exciting the glycogenolytic process—is investigated by producing disturbances involving various of the ductless glands, particularly the pancreas, the adrenals, the parathyroids and the pituitary. The influence of certain of these glands may be closely bound up with that exercised through the nervous control, as we have seen to be the case with the adrenal gland. Whether it is by the production of hormones directly necessary for proper carbohydrate metabolism, or by the removal from the blood of such substances as interfere with this process, that the ductless glands functionate, is one of the main problems we have to consider.

Utilization of Glucose in Tissues.—Although the experimental diabetes induced by disturbances in the function of the ductless glands is at first dependent on an upset of the glycogenic function and later of gluconeogenesis, the utilization of glucose in the tissues ultimately becomes interfered with. It is therefore important that we should digress for a moment to consider briefly what is known regarding the process by which sugar becomes utilized in the organism. That glucose becomes used up by active muscle there can be no doubt. Thus, if the muscles of one leg in the frog are tetanized, the glycogen content, compared with that of the other leg, will be found to be diminished.

At first sight it might appear that the easiest way to study the utilization of glucose in the muscles would be to compare its concentrations in the blood flowing to and coming from the muscle. The muscle that has been most successfully employed in studies of this kind has been the heart. Some years ago Starling and Knowlton[24] examined the consumption of sugar by the excised mammalian heart, and in their earlier experiments seemed to be able to show that the extent to which this consumption occurred was 4 milligrams per 100 grams heart muscle per hour. A more thorough repetition of these experiments later by Patterson and Starling[25] showed, however, that the results can furnish no criterion of the actual consumption of glucose by the tissue on account of the fact that the tissue itself may store away large quantities of carbohydrate in an unused state—i. e., as glycogen.

Other investigators have thought to study the utilization of glucose by observing the rate at which it disappears from drawn blood kept in a sterile condition at body temperature for some hours after death. This process is called *glycolysis,* and it has been assumed that the process is similar to that which occurs in the tissues themselves—an assumption, however, for which there is no warranty. Indeed, it may readily be shown that the glycolysis occurring in blood has very little if anything to do with the utilization of sugar in the tissues, for it has been found that glucose disappears from drawn blood very slowly indeed when

compared with the rate at which it disappears from the blood of animals in which the addition of glucose from the liver has been prevented by removal of this viscus (Macleod).[26]

A third method for studying the utilization of glucose consists in observing the *respiratory exchange* of animals. In normal animals the injection of glucose causes an increase in the carbon-dioxide excretion and a rise in the respiratory quotient, which it will be remembered is a ratio expressing the relationship between the amount of carbon dioxide excreted and of the oxygen retained in the organism. When carbohydrate is undergoing combustion, the quotient is nearly 1, whereas with that of protein it is about 0.7 (see page 547). By observing the quotient under given conditions one can compute the proportions of carbohydrate and of fat and protein that are undergoing metabolism. In the hands of Murlin and others,[27] this method has proved of some value in settling certain questions concerning the utilization of glucose in normal and diabetic animals; but the results must be interpreted with great care on account of the fact that temporary changes in the blood may cause a greater or a less expulsion of carbon dioxide from it. Thus, if acids appear in the blood, they will dislodge carbon dioxide from it, and apparently cause the respiratory quotient to rise. Alkalies, on the other hand, apparently cause it temporarily to fall, and unless the observations are done over a long period of time and with great care, faulty conclusions are very apt to be drawn from the results.

Diabetes and the Ductless Glands

We are now in a position to consider the forms of experimental diabetes produced by disturbances in the ductless glands.

Relationship of the Pancreas to Sugar Metabolism.—In no other of the many causes of diabetes has greater interest been shown than in that due to disturbance in the pancreatic function. Many of the earlier clinicians that followed cases of diabetes mellitus into the postmortem room, noted that definite morbid changes in the pancreas were a frequent accompaniment of the disease. Prompted by these observations, several investigators attempted experimental extirpation of the gland, but did not succeed in producing glycosuria in the few animals that survived the operation. Their failure, no doubt, resulted from incomplete extirpation. To reduce the severity of the operation, Claude Bernard injected oil into the pancreatic duct, and tied it; but he succeeded in keeping only two dogs alive for any length of time, and these did not exhibit glycosuria. Neither were other investigators that adopted similar methods any more successful. It looked as if the pancreas had very little to do with the cause of diabetes. In the year 1889 Minkowski

and von Mering in Germany, and de Dominicis in Italy, by thorough extirpation of the gland, succeeded in producing in dogs a marked and persistent glycosuria, accompanied by many of the other symptoms of diabetes. The first two authors attributed the condition to removal of an internal secretion.

The course of the diabetes thus produced is, however, somewhat different from that usually observed in man. It is extremely acute from the start, the G: N ratio being 1:3.6 (see page 664), and it is unaccompanied by any of the classical symptoms seen in the clinical condition. Experimental pancreatic diabetes can, however, be made to simulate very closely the disease in man. This was first of all demonstrated by Sandemeyer, who found that if the greater part of the pancreas was removed, the animals for some months, if at all, were only occasionally glycosuric, but later became more and more frequently so, until at last the condition typical of complete pancreatectomy supervened. Similar results have more recently been obtained by Thiroloix and Jacob, in France, and by Allen in this country. These investigators point out that different results are to be expected according to whether the portion of pancreas which is left does, or does not, remain in connection with the duodenal duct. When this duct is ligated, atrophy of any remnant of pancreas that is left is bound to occur, and this is associated with rapid emaciation of the animal, diabetes and death. When the remnant surrounds a still patent duct, a condition much more closely simulating diabetes in man is likely to become developed—one, namely, in which there is, for some months following the operation, a more or less mild diabetes, which, however, usually passes later into the fatal type.

It is, of course, difficult to state accurately what proportion of the pancreas must be left in order that the above described condition may supervene. Leaving a remnant amounting to from one-fifth to one-eighth of the entire gland is commonly followed by a mild diabetes, whereas if only one-ninth or less is left, a rapidly fatal type develops. As in clinical experience, the distinguishing feature between the mild and the severe types of experimental pancreatic diabetes is the tolerance toward carbohydrates. In the mild form, no glycosuria develops unless carbohydrate food is taken; in the severe form, it is present when the diet is composed entirely of flesh. It is thus shown that "by removal of a suitable proportion of the pancreas, it is possible to bring an animal to the verge of diabetes, yet to know that the animal will never of itself become diabetic. . . . Such animals, therefore, constitute valuable test objects for judging the effects of various agencies with respect to diabetes"—(Allen[18]). It therefore becomes theoretically possible to investigate, on the one hand, other conditions which will have an influence

even of burning the glucose which is derived from this foodstuff. By persistence in the starvation, however, the ability of the organism to utilize carbohydrate usually becomes so far restored that enough burns to prevent acidosis. Every case of diabetes can not, therefore, be expected to react in the same way to starvation, the determining condition being the relation between the quantities of glycogen and fat stored in the body at the outset of the fasting period. This relationship depends on the nature of the previous diet.

To sum up, "fasting will lower acidosis either in health or in diabetes, if it has the effect of stopping a one-sided metabolism and throwing the tissues on a more nearly balanced ration of fatty acids and glucose"—(Woodyatt). A practical point may be noted here—namely, that there is likely to be more danger of serious acidosis developing during starvation in fat than in lean diabetics. The importance of our appreciation of these facts in the starvation treatment of diabetes will be self-evident.

It must not be imagined, however, that no change takes place in the fat during its absorption and before it becomes deposited in the tissues. On the contrary, the stamp of individuality is put upon the fat, for, as we have already seen, its iodine value may become altered and its melting point changed during the process of absorption. In other words, although the absorbed fat does not become entirely adapted to conform with the ordinary qualities of the depot fat, yet it tends to change in this direction.

That some of the depot fat comes from carbohydrate is well known to stock raisers. When, for example, an animal is fed on large quantities of carbohydrate and kept without doing muscular exercise, its tissues become loaded with fat. If we desire strict scientific proof for this, we do not need to go further than the old experiments of Lawes and Gilbert, who, it will be remembered, showed that the fat deposited in the tissues of a growing pig is greatly in excess of the fat that could have been derived from the fat or protein which was meanwhile metabolized. The experiment was performed on two young pigs from the same litter and of approximately equal weight; one was killed and the exact amounts of fat and nitrogen in the body determined; the other was fed for several months on a diet the fat and protein contents of which were accurately ascertained. When after four months this pig was killed and the fat determined, it was found that much more had become deposited than could be accounted for by the fat and protein of the food, even supposing that all the available carbon of the protein had become converted into fat. The only conclusion is that the carbohydrate must have been an important source of the extra fat.

The Destination of the Depot Fat.—The depot fat becomes mobilized and transported by the blood to the active tissues whenever the energy requirements of the body demand it. During starvation, as we have seen, the depot fat is used to supply 90 per cent of the energy on which the animal maintains its existence. Before the fat is transported, it is probably broken down into fatty acid and glycerine, as which it passes through the cell walls to be again reconstructed into neutral fat in the blood. What agency effects this constant breakdown and resynthesis of fat it is difficult to say. Two ester-splitting enzymes are present in blood, one acting mainly on simple esters, the other on glycerides; but it has been impossible to demonstrate any evident relationship between either of them and the extent of fat mobilization.

The Fat in the Liver.—The physiology of the liver fat has been very diligently studied, particularly by Leathes and his pupils.[50] The outcome of this work has been to show that the liver occupies an extremely important position in the metabolism of fat, being, as it were, the half-

way house in the preparation of the fatty-acid molecule for consumption in the tissues. Fat is a material containing large quantities of potential energy. While in the depots this potential energy is so locked away as to be unavailable for tissue use. To make it available the depot fat is carried to the liver, where the energy becomes unlocked but not actually liberated. The fat is then transported to the tissues, and the liberation of the energy occurs. Neutral fat is like wet gunpowder: it contains much potential energy, but not in a suitable condition for explosion. The liver, as it were, dries this gunpowder, whence it is sent to the tissues to be exploded.

The great importance of the liver in fat metabolism is indicated by comparison of the percentages of fat—or better of fatty acid—contained in it under different conditions of nutrition. In the ordinary run of slaughter-house animals the liver contains from 2 to 4 per cent of higher fatty acid, but in about one in every eight animals a much higher percentage will be found to occur. The same is true in laboratory animals. In the case of the human liver as obtained from autopsies in certain classes of patients, from 60 to 70 per cent of the dry weight of the organ, or 23 per cent of the moist weight, may be fatty acid. There is no other organ in the animal body that is ever loaded with fat to this extent. As in the depots, this liver fat might be derived either from fat carried to the organ from elsewhere in the body, or it might represent a surplus of manufactured fat.

That *transportation of fat to the liver* occurs is very readily demonstrable both in the laboratory and in the clinic. About forty hours after giving phlorhizin to a dog, it has been found that enormous quantities of fat appear in the liver; a few hours later, however, this excess of fat may have entirely disappeared. Fatty infiltration of the liver is also observed in phosphorus poisoning, although in this case the fat usually persists till the death of the animal. In man, in delayed chloroform poisoning and in cyclical vomiting, enormous quantities of fat may be present in the liver within a very short period of time after the onset of the condition. There can therefore be no doubt that fat is transported to the liver under abnormal conditions, but this can not be taken as evidence that the liver has anything to do with fat metabolism in the normal animal. Such evidence has been supplied by Coope and Mottram,[51] who have been able to show that, at least in rabbits, a similar invasion of the liver with fat occurs in late pregnancy and early lactation. They also found that the fatty acid deposited in the liver in late pregnancy gives an iodine value which lies nearer to that of the mesenteric fatty acid than is the case in normal animals. Mottram concludes that "wherever . . . there is abundant fat metabolism, the

liver is found to be infiltrated with fats, presumably to be handed on elsewhere when worked up." It is interesting that the fetus is greedy of unsaturated fatty acids.

The most likely *source of the fat transported to the liver* is the fat present in the depots, unless when digestion is in progress, when it may be the fat from the intestine. That much of it comes from the *depots* is easily demonstrated. Thus, the more extensive the infiltration of the liver with fat, the more closely will this fat be found to agree with the depot fat in its chemical characteristics. This has been very clearly shown by, first of all, starving an animal so as to clear the depots of fat as much as possible; then feeding it on some "ear-marked" fat (unusual melting-point or a brominized fat); and after another day or so of starvation, so as to clear the liver itself of fat, poisoning the animal with phosphorus or phlorhizin. The liver will be found shortly afterwards to be invaded with fat which has all the ear-marks of that on which the animal had been fed.

Evidence of the same character has been furnished in a series of clinical cases by observations on the amount of fat and the iodine value of the fatty acid of the liver. This is shown in the accompanying table.

FATTY ACIDS OF LIVER

	Cause of Death	Higher Fatty Acids per cent of Dry Weight	Iodine Value of Fatty Acids
Normal figures	1. Pernicious anemia	12.1	116.8
	2. Lobar pneumonia	13.7	116.8
	3. Pernicious anemia	14.25	116.0
	4. Diabetes	14.4	119.6
Commencing fatty change	5. Toxemic jaundice	15.6	109.5
	6. Accident	17.2	103.5
	7. Empyema	21.5	96.0
	8. Phthisis	25.4	96.4
Marked fatty change	9. Broncho-pneumonia	38.4	84.9
	10. Appendicitis	44.0	91.1
	11. Carcinoma of bladder	47.2	77.8
	12. Broncho-pneumonia	54.6	71.8
	13. Ulcerative colitis	60.9	80.3
	14. Accident	66.3	63.0
	15. Dysentery	73.5	69.1

This table clearly shows that the more fat there is in the liver, the nearer this fat approaches in character that stored in the depots.

That some of the fat in the liver may come directly *from the fat recently absorbed from the intestine* is also very readily demonstrable. Thus, when cocoanut oil was placed in the intestine of anesthetized animals, along with bile salts and glycerine, it was found by Raper[52] that 30 per cent of the absorbed oil appeared in the liver.

That higher fatty acids, such as caproic ($C_6H_{12}O_2$) and caprylic ($C_8H_{16}O_2$), have actually been isolated from the products of this fermentation is a very significant fact, and it is of interest to note that Leathes has sometimes found an increase in higher fatty acids to occur during the aseptic incubation of liver pulp. Unfortunately, however, the increase of fatty acid could not be shown to be affected by adding substances to the liver which, according to the above equations, should yield fatty acid.

The Method by Which the Fatty Acid is Broken Down.—In the chemical laboratory, ordinary oxidizing agents attack the fatty-acid chain at the C-atom next the carboxyl (COOH) group (the alpha C-atom). But this can not occur in the animal body, because it would leave behind a smaller chain containing an uneven number of C-atoms, and such chains are never found present in the animal fats. On the contrary, the commoner fats all contain an even number of C-atoms, thus: Butyric, $C_4H_8O_2$; palmitic, $C_{16}H_{32}O_2$; stearic, $C_{18}H_{36}O_2$; oleic, $C_{18}H_{34}O_2$.

The intermediary substances which are produced during the gradual breakdown of the fatty-acid molecule in the normal animal are of a very transitory character so much so indeed that it is impossible for any one of them to accumulate in sufficient amount to permit of isolation, or even detection, by chemical means. How then are we to *identify the intermediary products?* This has been rendered possible by the discovery that, when anything occurs to disturb the normal course of fat metabolism, as, for example, when the tissues are deprived of carbohydrates (as in starvation or in severe diabetes), the oxidation of the fatty-acid chain stops short when a chain of four C-atoms still remains unbroken. These last four C-atoms seem to form a residue that is more resistant to oxidation than the remainder of the fatty-acid molecule. It is a residue, therefore, which is quite readily further oxidized to CO_2 and H_2O under normal conditions, but which, although incapable of becoming completely oxidized when the metabolism is upset, does undergo a partial oxidation, resulting in the production of various intermediary products. These accumulate in the body in sufficient amount to overflow into the urine, from which they can be isolated and identified.

The fatty acid with 4 C-atoms is *butyric,* $CH_3CH_2CH_2COOH$, and the first oxidation product formed from it in the body seems to be *β-oxybutyric acid,* $CH_3CHOHCH_2COOH$. This then becomes oxidized to form a body having the formula CH_3COCH_2COOH, *acetoacetic acid,* which, on further oxidation, readily yields CH_3COCH_3, or *acetone.* These substances (β-oxybutyric acid, acetoacetic acid and acetone) appear in the urine during carbohydrate starvation, as in diabetes.

It might be objected, however, that a chemical process occurring under abnormal conditions need not also occur in the *normal* animal. That it probably does, however, is indicated by the results of the experiments

of Knoop and of Embden and his coworkers. Knoop conceived the idea of introducing into the fatty-acid molecule some group which is resistant to oxidation in the body. The phenyl group (C_6H_5) was found to have this effect. By feeding an animal with the phenyl derivatives of acetic, propionic, butyric, and valeric acids, it was found that the urine contained either hippuric (see page 630) or phenaceturic acid. Both of these are compounds of aromatic acids with glycocoll or aminoacetic acid (CH_2NH_2COOH), one of the protein building-stones and always available in the organism to form such compounds, thus:

$$(1)\quad \underset{\text{(benzoic acid)}}{C_6H_5COOH} + \underset{\text{(glycocoll)}}{CH_2NH_2COOH} = \underset{\text{(hippuric acid)}}{C_6H_5CONHCH_2\ COOH.}$$

$$(2)\quad \underset{\text{(phenylacetic acid)}}{C_6H_5CH_2COOH} + \underset{\text{(glycocoll)}}{CH_2NH_2COOH} = \underset{\text{(phenaceturic acid)}}{C_6H_5CH_2CONHCH_2COOH.}$$

When either benzoic acid (C_6H_5COOH) or phenylacetic acid ($C_6H_5CH_2$-$COOH$) is formed in the body as a result of the oxidation of phenyl derivatives of the higher fatty acids, the acid combines with glycocoll according to the above equations. From this it follows that if oxidation occurs so that two C-atoms are thrown off at a time (β-oxidation), fatty acids with an even C-atom chain should yield hippuric acid, and those with an uneven chain, phenaceturic. This was found to be the case, as the accompanying table shows.

ACID FED	OXIDATION PRODUCT	EXCRETED AS
Benzoic acid, $C_6H_5.COOH$	Not oxidized	Hippuric acid
Phenylacetic acid, $C_6H_5.CH_2.COOH$	Not oxidized	Phenaceturic acid
Phenylpropionic acid, $C_6H_5.CH_2.CH_2.COOH$	$C_6H_5.COOH$	Hippuric acid
Phenylbutyric acid, $C_6H_5.CH_2.CH_2.CH_2.COOH$	$C_6H_5.CH_2.COOH$	Phenaceturic acid
Phenylvaleric acid, $C_6H_5.CH_2.CH_2.CH.CH_2.COOH$	$C_6H_5.COOH$	Hippuric acid

(From Dakin.)

Embden's experiments are equally convincing. He studied the formation of acetone in defibrinated blood perfused through the freshly excised liver. Normally only a trace of this substance is formed, but when fatty acids with an even number of carbon atoms were added to the blood, they gave rise to a marked increase in acetone, whereas those with an uneven chain failed to cause any change. The acetone was found to be derived immediately from acetoacetic acid. The following table shows the results.

	NORMAL FATTY ACID	FORMATION OF ACETOACETIC ACID
Acetic acid	$CH_3.COOH$	–
Propionic acid	$CH_3.CH_2.COOH$	–
Butyric acid	$CH_3.CH_2.CH_2.COOH$	+
Valeric acid	$CH_3.CH_2.CH_2.CH_2.COOH$	–
Caproic acid	$CH_3.CH_2.CH_2.CH_2.CH_2.COOH$	+
Heptylic acid	$CH_3.CH_2.CH_2.CH_2.CH_2.CH_2.COOH$	–
Octoic acid	$CH_3.CH_2.CH_2.CH_2.CH_2.CH_2.CH_2.COOH$	+
Nonoic acid	$CH_3.CH_2.CH_2.CH_2.CH_2.CH_2.CH_2.CH_2.COOH$	–
Decoic acid	$CH_3.CH_2.CH_2.CH_2.CH_2.CH_2.CH_2.CH_2.CH_2.CH_2.COOH$	+

(From Dakin.)

For a long time it was difficult for chemists to understand how such a process of oxidation at the β-C-atom could occur, since they were unable to bring it about in the laboratory by the use of the ordinary oxidizing agents, but recently Dakin has removed the difficulty by showing that hydrogen peroxide (H_2O_2) oxidizes fatty acids just exactly in this way.

We may sum up the results of these experiments and observations by stating that *normal saturated fatty acids and their phenyl derivatives can undergo oxidation, not only in the animal body, but also in vitro, in such a manner that the two (or some multiple thereof) termial C-atoms are removed at each successive step in their decomposition.*

But we must not be too hasty in concluding from these experiments that the steps in the process are *necessarily* in the order of first, the production of a β-hydroxy acid, and second, the oxidation of this to a ketone group. The mere presence, side by side, of β-hydroxybutyric acid and of acetone in the above experiments does not indicate which is the antecedent of the other, and indeed there are several experimental facts that seem to show that the hydroxy acid may be derived from the ketone. For example, when acetoacetic acid is added to minced liver and the mixture incubated, β-hydroxybutyric acid is formed (a reduction process), although less usually the reverse action (oxidation) may occur when β-hydroxy acid is added. A reversible reaction must therefore be capable of occurring between these two substances, thus:

$$\underset{(\beta\text{-oxybutyric acid})}{CH_3.CHOH.CH_2.COOH} \;\underset{\text{oxidation} \;\longrightarrow}{\xleftarrow{\text{reduction}}}\; \underset{(\text{acetoacetic acid})}{CH_3.CO.CH_2.COOH.}$$

We know practically nothing as to the conditions determining whether oxidation or reduction shall predominate, but there are two significant facts that one should bear in mind: (1) that a plentiful supply of oxygen is necessary for the oxidative process, and (2) that the presence of readily oxidizable material in the liver (e.g., carbohydrates) may determine the direction which the reaction shall take. It is commonly said that fats burn in the fire of carbohydrates, and it may be that the un-

doubted diminution in acidosis which occurs in diabetes when carbohydrate food is given is dependent upon the directive influence which its combustion in the liver has on the above processes. But we must be cautious not to transfer results obtained by experiments with minced liver in judging of the reactions which occur during life. Provisionally, then, we must assume either that β-hydroxybutyric acid is a necessary stage in the oxidation of butyric acid or that it is formed by reduction of acetoacetic acid, which is really the first step in that process.

Of course there is no evidence in the above experiments that the higher fatty acids are also broken down by the removal of two C-atoms at a time, nor has it been possible to detect any ketonic or β-hydroxy derivatives of them in the animal body. We can only reason from analogy that a similar process may occur, although some support is furnished for such a view by the fact that ketonic fatty acids have been found in vegetable organisms.

What, then, it may be asked, is the relation of the desaturation of fatty acids which we have seen occurs in the liver (and probably elsewhere) to the β oxidation? There can be no doubt that both processes can occur in the animal body, indeed in the same organ, e.g., the liver; and it is important to ascertain their relationship to each other. The conclusion which would seem to conform best with the known facts is that the desaturation process occurs (in the liver) so as to break up the long fatty-acid chain into smaller chains, which are then capable of β oxidation (in the tissues); desaturation may be the process by which the molecule is rough hewn, and β oxidation that by which the resulting pieces are finally split to their smallest pieces—that is, to molecules of the size of acetic acid, *which are finally* completely burnt to carbonic acid and water.

The increase of iodine value observed by Leathes and his coworkers need not, as has already been pointed out, necessarily indicate that new double linkages have been introduced in the fatty-acid chain; it may merely indicate that structurally isomeric derivatives which absorb iodine more readily have been formed. Direct evidence of desaturation has, however, been offered by Hartley, who isolated the unsaturated fatty acids (by dissolving the lead soaps in ether) from pig's liver and then proceeded to oxidize them with alkaline permanganate. When the olein of the depot fat is thus treated at a low temperature, two hydroxyl groups become attached where the double linkage existed (forming dioxystearic acid), and when the mixture is now warmed, the molecule splits into two at this place, forming two lower acids (pelargonic and azelaic):

$$(1)\quad \underset{\text{(oleic acid)}}{CH_3-(CH_2)_7CH:CH(CH_2)_7COOH};$$

$$(2)\quad \underset{\text{(dioxystearic acid)}}{CH_3-(CH_2)_7-\overset{OH}{CH}\text{———}\vdots\text{—}\overset{OH}{CH}\text{——}(CH_2)_7COOH};$$

$$(3)\quad \underset{\text{(pelargonic acid)}}{CH_3\ (CH_2)_7COOH} + \underset{\text{(azelaic acid)}}{COOH-(CH_2)_7COOH}.$$

We may conclude from this that the double linkage in the oleic acid of the depot fat exists between the ninth and tenth C-atoms. But it is otherwise in the case of the unsaturated acid from the liver (pig's), for under the above process of oxidation this yielded caproic acid, which, since this acid has six C-atoms, would indicate that the double linkage existed between the sixth and seventh C-atoms. Another interesting fact brought to light by the experiments was that a tetraoxystearic acid was formed, which fell apart in such a way as to indicate that the hydroxyl groups occurred between the sixth and seventh and between the ninth and tenth C-atoms. The occurrence of this substance would be satisfactorily explained by the introduction into the molecule of oleic acid of a second double bond—i. e., between the sixth and seventh C-atoms. "The acids found in the pig's liver may be accounted for, in other words, by supposing that desaturation of stearic acid and of the ordinary (depot) oleic acid occurs by the introduction of a double link between the sixth and seventh carbon atoms in each case"—(Leathes). Still other double links may, however, be introduced into the fatty-acid chain, for acids of the linolic acid series are present in cod-liver oil. Finally, it is of interest to note that caproic acid is a product of the above oxidation process, for it has an even number of C-atoms and therefore will form β-oxybutyric acid.

To go into these chemical problems any further here would be out of place. One other fact, should, however, be borne in mind—namely, that the unsaturated acids may be formed from saturated acids through the intermediate formation of β-hydroxy and β-ketonic acids. Their mere presence, in other words, should not be taken as evidence that the oxidation of fatty acids is initiated by the introduction of an hydroxyl group at the β position in the chain.

FEVER

The clinical application of a knowledge of the mechanism of heat regulation in the animal body concerns the causes of fever. In the most familiar form fever is produced by infectious processes, but it may also be owing to various other causes, among which may be mentioned the parenteral injection of foreign protein, excessive destruction of protein substances in the body itself, the action of certain drugs, and lastly, injury to the base of the brain or lesions of the upper levels of the spinal cord. Various types of fever are recognized: when the temperature remains constantly above the normal, it is known as continuous fever; when oscillations occur but the temperature never falls to the normal level, it is known as remittent; when it attains the normal level at certain periods during the day, it is known as intermittent.

Causes of Fever

During a *sudden rise* in temperature there is, on the one hand, increased heat production in the muscles, and on the other, diminished heat loss from the surface of the body. *The fever is therefore due to an exaggeration of the processes by which the body normally reacts to conditions which tend to lower the body temperature.* The increased muscular activity thus induced often causes visible contractions, familiar as shivering; and the constriction of the cutaneous blood vessels produces the subjective sensation of chills, and causes the skin to become pale and cold to the touch. The skin muscles also contract, producing "goose skin." During this stage, objective demonstration of the curtailment of the skin circulation can be secured by observation of the bloodflow through the hands and feet (page 283). When the temperature *suddenly falls* again, the crisis, as it is called, muscles become flaccid and produce less heat, and the cutaneous blood vessels dilate, as has been shown by measurements of the bloodflow of the hands and feet. At the same time also, the sweat glands are stimulated and marked perspiration occurs.

Concerning *the cause of continuous fever,* it must be assumed that the balance between heat production and heat loss has been adjusted at a higher plane than normal. We can not explain the fever on the basis either that heat production is permanently increased or that heat loss is permanently diminished, for in neither of these cases would the temperature stand at a permanent level but would steadily rise or fall, according to which mechanism was disturbed. While set at this higher plane of fever, the thermogenic nerve centers are still capable of responding in the usual way to the influences which cause the body tem-

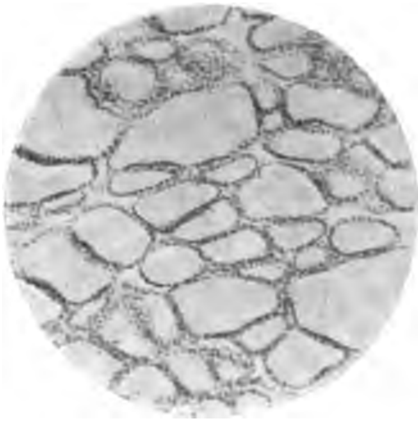
hypersecretion of epinephrine into the blood. The results seemed to confirm the hypothesis. In the first place, it was found that, whereas the blood drawn from the vena cava opposite the entry of the adrenal veins (by passing a catheter up the femoral vein till its free end lay at this level) in a normal male cat did not give evidence of the presence of epinephrine when tested by means of the intestinal segment method, it did so in a cat that had previously been frightened by allowing a dog to bark at it. Such results were not obtained after removal of the adrenal gland, or in a female cat, which is usually indifferent to such a method of frightening. Cannon also thought that many of the other adaptations which take place in an animal in this condition could be attributed to the presence of an excess of epinephrine in the blood. The three most important of these are: (1) increased discharge of sugar from the liver into the blood; (2) increased efficiency of muscular contraction; (3) diminished clotting time of the blood—all of which are adaptations enabling the animal either to conquer the source of the fear or to be in a better position to recover from any bodily injury involving a loss of blood should he suffer bodily damage. Stewart and Rogoff have more recently thrown considerable doubt on these conclusions by finding that cats in which both adrenal glands are entirely removed from the influence of the nervous system, behave like normal animals when frightened, and develop hyperglycemia when asphyxiated or etherized. It is scarcely necessary to point out that, until it is definitely established by experimental investigation that epinephrine may be discharged in excessive amounts under certain conditions, it is irrational to assume that such may occur in disease. *The surgical removal of the adrenal gland is certainly not warranted under any circumstances.*

The Association of the Adrenal with Other Endocrine Organs

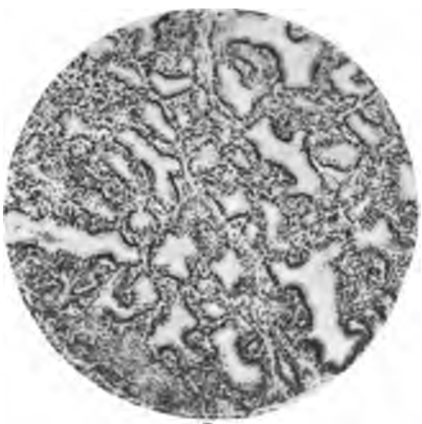
We have at present very little accurate and reliable information on the association of the adrenal with other endocrine organs. That epinephrine has an influence on many diverse organs and glands is an undoubted fact, but this is more probably to be attributed to an activating influence on sympathetic nerve endings than to any specific relationship between the adrenal glands and the gland in question. The most important of the results that have been obtained are the following:

1. **With the Thyroid and Parathyroid.**—Cannon and Cattell, after confirming Bradford's discovery that an electric current of action is set up in the salivary gland when it is excited to activtiy, proceeded to investigate the occurrence of such a current in the thyroid gland.[73] By placing one nonpolarizable electrode on the gland itself and the other on the neighboring subcutaneous tissues or on the trachea, a current was found to be set up by stimulation of the sympathetic nerve supply of the thyroid, by intravenous injection of epinephrine, or by stimulation of the great

and it can be produced in normal animals (particularly rats) by placing them on an excessive meat diet. Important observations bearing on this point have been made by Marine on brook trout, in which it has been found that the so-called carcinoma that develops when the fish kept in hatcheries are fed with unsuitable food and overcrowded, is really a typical hyperplasia. In its second stage this develops into what is known



A.



B.

C.

Fig. 193.—Microphotographs of thyroid gland of dog. *A*, normal hyperplasia; *B*, active hyperplasia; *C*, colloid goiter. (From Marine and Lenhart.)

as ***colloid goiter*** which is produced by a deposition of colloid material between the rows of cells so as to cause an opening out again of the vesicles (Fig. 193), with a consequent tendency to a reversion to the normal histologic structure, so far as this is possible. The vesicles in such a gland are of enormous size, and the lining epithelium, low cubical, or almost flat in shape.

The outstanding characteristic feature of the colloid material is that

it contains iodine, which exists in combination with a nonprotein nitrogenous base, and is usually called iodothyrin. In the gland itself the iodothyrin may be in combination with protein, forming iodothyroglobulin. E. C. Kendall[79] has recently succeeded in isolating a pure crystalline substance of perfectly constant composition and containing over 60 per cent of iodine. It is called *alpha-iodine,* and has been identified as an indole compound. In extremely minute dosage it greatly affects the energy metabolism, and is said to induce symptoms like exophthalmic goiter. Its therapeutic value in cases of thyroid deficiency is remarkable. Kendall believes this substance to be the active constituent of the thyroid and to be associated with the metabolism of amino acids. For one thing, when it is given alone no change occurs in pulse rate, whereas if amino acids are given along with it, there is acceleration.

The importance of the relationship between the function of the thyroid and the iodine-containing material is indicated by the changes which occur in the percentage of iodine in the glands under varying conditions of activity. Marine observed that the amount of iodine is inversely proportional to the degree of hyperplasia of the gland, and when the hyperplastic condition becomes fully developed, scarcely a trace of iodine is contained in the gland. Later, when the hyperplasia gives place to colloid goiter, the iodine increases again, both absolutely and relatively. Moreover, it has been found that if iodide be administered to an animal suffering from hyperplasia, the hyperplastic condition very quickly disappears (Fig. 192) and the animal becomes normal. Thus, in brook trout, the poor nutritive condition of the fish when hyperplasia has developed can be immediately remedied by placing them in larger quantities of running water or by adding small traces of iodide to the water. The administration of small amounts of iodine as in ordinary salt from salt deposits also prevents goiter in farm stock, this having been first noted in the State of Michigan, where prior to the discovery of salt deposits sheep breeding was an entire failure. The importance of administering small doses of iodides to school children living in goitrous districts has recently been emphasized by Marine and Kimball.[78] As small a dose as 0.001 gm. at weekly intervals prevents goiter in puppies susceptible to it.

Experimental Thyroidectomy

A correct interpretation of the functional changes and symptoms which follow upon partial or complete removal of the thyroid gland, or from its disease, has proved a very difficult problem, partly because sufficient care has not been taken to note how much parathyroid tissue was removed along with the thyroid, and partly because the fact has been over-

looked that the effects produced by thyroidectomy and parathyroidectomy are often very different in animals of the same kind at different ages. Speaking generally, it may be said that the influence of the parathyroid is focused mainly on the nerve centers and only to a secondary degree on the metabolic functions, whereas the reverse is the case with the thyroid, its main effect being on metabolism, although it probably also exercises a secondary effect on the nerve centers. More so than in the case of any other endocrine organ, our knowledge concerning the function of the thyroid has been gained by clinical experience, and it is difficult to say whether the clinical or the experimental method has contributed the greater amount of information.

The results of experimental extirpation of the thyroid vary according to the age of the animal, and frequently they are by no means marked, provided sufficient parathyroid tissue has been undamaged. The symptoms are in general thickening and drying of the skin, with a tendency to adiposity and a loss of tone of the muscle. The body temperature is low and the sexual functions become subnormal. Nervous symptoms in the direction of mental dullness and lethargy are also usually present. Surgical removal of the thyroid in man produces the condition known as *cachexia strumipriva.* The symptoms may first of all become apparent a few days after the operation, or they may remain latent for years, and then develop so as to produce the condition known as myxedema. When nervous symptoms are prominent in cachexia strumipriva, it is usually taken as evidence that an excessive amount of parathyroid tissue has been destroyed. Kocher states that after complete loss of the thyroid, life is impossible for more than seven years, and that to prevent ultimate ill effects, at least one-fourth of the organ should be left intact.

Disease of the Thyroid

The symptoms of diseased conditions of the thyroid may be interpreted as the consequence of increased or diminished functioning of the gland. Sometimes, however, the less active gland is really increased in bulk, this increase being caused by the accumulation in it of very large quantities of colloid material accompanied by an attenuated condition of the vesicular cells (see page 751). When the gland is atrophied at birth, the condition of *cretinism* soon becomes developed (Fig. 194). The characteristic features of cretinism are: (1) An arrest of growth, especially of the skeleton, accompanied by incomplete ossification of the long bones and failure of the fontanelles of the skull to close properly. (2) Poor development of the muscular system. (3) An unhealthy, dry, swollen condition of the skin, so that it is yellowish in color, the face being pale

and puffy. (4) An abnormal development of the connective tissues causing a shapeless condition of the surface; the abdomen is always swollen, the hands and feet are shapeless, and the nose depressed. (5) The nervous system also fails to develop properly, so that at the age of puberty or over, the child remains like an infant in his mental behavior, idiotism being common. Indeed, the whole clinical picture is so characteristic that once having seen a case no one can fail afterward to

Fig. 194.—Cretin, nineteen years old. The treatment with thyroid extract started too late to be ot benefit. (Patient of Dr. S. J. Webster.)

recognize the disease. Besides being due to congenital absence of the thyroid (sporadic type), cretinism may also occur as a result of goitrous degeneration of the gland. This forms the so-called endemic variety of the disease, and is more commonly seen in goitrous districts, being not infrequently associated with disease of the parathyroid, in which case the nervous symptoms are very prominent.

Atrophy of the thyroid in adults causes the clinical condition known

as *myxedema,* and here again the symptoms are very characteristic (Fig. 195). The skin is dry and thick, with a deposition of connective tissue often containing fat in its deeper layers; the hands and feet become unshapely; the lips thick and the tongue somewhat enlarged, so that when the person attempts to speak, it appears as if the tongue were too large for the mouth; the hair falls out; there is a low body temperature, and it can be shown that the energy metabolism is greatly depressed, and that a deficiency of oxygen is being consumed. It is said the person can take a larger quantity of sugar than an ordinary individual without the development of glycosuria, but the depression of the metabolic function causes the patient to take sparingly of food, in spite of which, however, the body weight may steadily increase. The sexual function becomes

A. *B.*

Fig. 195.—*A,* Case of myxedema; *B,* Same after seven months' treatment. (From Tigerstedt.)

depressed, and there is involvement of the nervous system as shown by mental dullness and lethargy.

Although the thyroid gland is much atrophied in myxedema, symptoms that are very similar may also occur when the gland is enormously enlarged. As already explained, however, this enlargement is due merely to an accumulation of colloidal material and is really an atrophic condition. A patient suffering from endemic goiter may at first exhibit symptoms which are usually attributed to a hypersecretion of thyroid material into the blood (the symptoms will be described immediately), but later these give place to symptoms not unlike those of myxedema.

It is concluded that the above conditions are due to deficiency of thyroid function, or *hypothyroidism,* because: (1) the gland is atrophied,

and (2) similar symptoms to those exhibited by the clinical conditions can be produced experimentally by the removal of the gland in animals. By observations on the effect of administration of thyroid extract to cretinous or myxedematous patients, prompt amelioration of the symptoms occurs, which certainly suggests that the real cause is the absence of an internal secretion. There is probably nothing more striking in the whole domain of therapeutics than this effect from the administration of thyroid extract or, more so still, of alpha-iodine. If the treatment is started early enough, the cretinous child from being an ill-developed idiot quickly catches up with children of his own age and becomes in every respect normal. Even if this treatment is not undertaken until the child is several years of age, it is remarkable how quickly the benefit may show itself. In myxedema and cachexia strumipriva also, the symptoms very quickly disappear and the person becomes perfectly normal by the treatment. In all these conditions, however, the thyroid extract must be administered continuously in order to prevent the reappearance of symptoms.

Quite distinct from the above described conditions of hypothyroidism are those produced by an excess of thyroid autacoid in the blood, namely, *hyperthyroidism.* Such a condition can be produced experimentally in normal animals by the administration of thyroid extract or alpha-iodine (Kendall). In man large doses are soon followed by great quickening of the pulse with some irregularity, flushing of the skin, increased perspiration, tremor in the limbs, emaciation, and marked nervous excitability. Along with these symptoms, metabolic investigations have shown that the energy output per square meter of surface is greatly increased, being sometimes nearly doubled; that the nitrogen excretion is excessive; and that alimentary glycosuria is very commonly present. The body temperature is not, however, as a rule increased, because although metabolism is excited, yet heat loss is correspondingly increased. Exophthalmos is said to develop very occasionally after such administration, but this is doubtful. Lastly, there are usually digestive disturbances, although the appetite is likely to be increased. The pulse is quickened after administration of alpha-iodine only when protein food is also taken. This is believed by Kendall to be due to the association between the thyroid hormone and the metabolism of the amino acids.

The symptoms following the injection of the extract are very similar to those of the disease known as *exophthalmic goiter*. Indeed, the symptoms are so much alike in the two conditions that it is scarcely necessary to describe them specially for the disease except to mention that the exophthalmos is much more likely to be present.

Like simple goiter this variety is from three to four times more fre-

quent in women than in men, a fact of significance when we recall the evidence of association between the thyroid gland and the generative organs. It is said that the disease is usually coupled with persistence of the thymus gland. The thyroid gland in exophthalmic goiter is enlarged, sometimes in one lobe; it is hard and pulpy, and on auscultation a murmur is heard. Histologically the gland presents a picture very like that which has been described above as hyperplasia; that is to say, the vesicles have a deficiency of colloid material; their epithelium is columnar and folded up into the vesicles; and the interstitial tissue between the vesicles is very markedly increased.

Exophthalmic goiter is almost universally claimed to be due to hypersecretion of the thyroid, because: (1) the symptoms of the disease are not unlike those produced by excessive administration of thyroid to a nórmal individual; and (2) they are in general opposite in character to the symptoms found in cases where the thyroid gland is atrophied. The blood of a person with exophthalmic goiter when injected into mice increases their resistance to the toxic action of acetonitrile, which is also the case after thyroid extract has been injected. In many cases of exophthalmic goiter partial removal of the gland is said to ameliorate the symptoms. Other clinicians, however, state that if the patient is given proper medical treatment, rest, and diet, equally beneficial results can be obtained.

Certain investigators, however, deny that it has yet been conclusively demonstrated that exophthalmic goiter is due to hypersecretion of the thyroid (Marine). It is pointed out that, if hypersecretion were the cause of the disease, one would expect that the injection into animals of the blood of patients suffering from it would produce symptoms similar to those following the injection of thyroid extract. The results of such experiments, however, have been extremely confusing and very indecisive, since it is difficult to recognize in laboratory animals many of the characteristic symptoms, especially those affecting the skin and eyes and the general bodily nutrition. Another difficulty in accepting the hypersecretion hypothesis is the fact that an extract of a gland removed from an exophthalmic patient has no different physiologic action on a normal animal from an extract of a normal gland containing the same percentage of iodine. The evidence is by no means conclusive one way or the other, and it may well be that the observed changes in the thyroid gland are not the cause of the symptoms of exophthalmic goiter, but merely, like the other symptoms of this disease, a result of some condition located elsewhere.

The Relationship of the Thyroid with Other Endocrine Organs

1. **With the Generative Organs.**—Evidence of an association between the female generative organs and the thyroid is very strong; thus, the

thyroid becomes enlarged at puberty, during the menses, and during pregnancy, and in thyroidectomized young animals the sexual glands fail to develop properly.

2. **With the Adrenal Glands.**—(See page 746.)

3. **With the Pituitary Body.**—After removal of the thyroid, the pituitary becomes greatly altered and enlarged, particularly the pars anterior, in which it is not uncommon to find that a certain amount of vesicles containing colloid, not unlike those of the thyroid, become developed. This colloid material, however, does not contain iodine. It is said that this increase of the pituitary after thyroidectomy does not occur if thyroid extract be administered. Increased activity of the pars intermedia of the pituitary is also quite plain. These facts would at first sight seem to indicate that the pituitary and the thyroid can act vicariously, but this is very doubtful, for it has not been found that pituitary extract has any beneficial effect in the treatment of goiter and myxedema. Nevertheless the association in function of the two glands must be more or less close, not alone for the above reasons, but also because they are both associated to much the same degree with the sexual organs, and both act on the higher functions of the nervous system in much the same manner.

4. **With the Thymus Gland.**—The persistence of the thymus in exophthalmic goiter, as well as the anatomic and embryologic relationship between thymus and thyroid, is taken to indicate some close relationship.

THE PARATHYROIDS

Experimental Parathyroidectomy

Experimental parathyroidectomy yields results which vary in different groups of animals, undoubtedly because of the fact that in some, such as the rat and rabbit, accessory parathyroids may exist. In general, however, it has been found that if more than two of the four parathyroids be removed, very definite and pronounced nervous symptoms soon supervene and if all four glands be removed, a quickly fatal result is inevitable. The most acute symptoms are exhibited by the carnivora. They may not be apparent for a day or two after the operation, although during the period the animal is in a depressed state, refusing food and losing weight rapidly. The muscles are also more or less stiff during this stage. When more definite symptoms appear, they consist of a marked abnormality of muscular contraction, leading to the occurrence of fibrillar contractions, or tremors and, later, to cramp-like and clonic contractions. When spontaneous movements are made, a

peculiar shaking of the foot, like that made by a normal animal to shake water off its pads, is a characteristic symptom. The slightest stimulation of the peripheral nerves is sufficient to induce one of these attacks, which recur with ever increasing frequency, becoming at the same time more pronounced and accompanied by other disturbances, such as diarrhea, profuse salivation, rapid pulse, and dyspnea (in the dog but not in the cat). In cases that are not quickly fatal, the hair tends to be shed, and the teeth to be improperly calcified (in young animals). Where a certain amount of parathyroid tissue has been left—for example, one of the four lobes—the symptoms may not appear except under conditions of special strain to the animal economy, such as pregnancy or improper diet. Thus, in a bitch from which three of the four glands had been removed, no symptoms of tetany occurred until she became pregnant. Under the same conditions it has been found that a diet of flesh is much more apt to bring about the condition than one of vegetables or milk.

Tetany, as the above condition is called, may also become developed in man either as the result of surgical removal of the parathyroids or because of their improper development. The symptoms in man are very similar to those observed in laboratory animals, the only difference being that the muscular contractions are more likely to be tonic in character. Certain symptoms that may develop during pregnancy or in the course of infectious diseases or in newborn infants have also been found to be associated with degeneration of or hemorrhage into the parathyroid (idiopathic tetany), and certain obscure nervous diseases in adults, such as paralysis agitans, may possibly also be associated with changes in this gland. Chorea, epilepsy, and eclampsia have likewise been thought to be associated with it.

The parathyroid gland, besides influencing the nerve centers, has also an influence on metabolism. The symptoms produced are: (1) rapid emaciation and failure to grow; (2) a tendency to the production of glycosuria, often detected by finding that the assimilation limit for carbohydrate is lowered (page 652); and (3) most definitely of all, an interference with calcium metabolism, as illustrated by the failure of the teeth and bones to calcify properly. This interference with normal metabolism led Kellogg and Voegtlin[81] to study the effect produced on parathyroidectomized animals by the administration of calcium. It was found that the symptoms were considerably ameliorated. These authors concluded from their results that the essential cause of tetany is a deficiency of calcium in the blood. It is possible however that the beneficial action of calcium salts in this condition is that it decreases the excitability of the nervous system, an action which it is known to possess.

When the tetany is the result of a complete extirpation of all parathyroid tissue, the symptoms can be combated by a successful transplantation or graft of parathyroid tissue made from an animal of the same species. Indeed, it has been found that the success of a graft of parathyroid is assured only when the graft is derived from the same kind of animal as that from which the parathyroid has been removed. Implantation into the subcutaneous tissue of a tetany patient of parathyroid tissue obtained fresh from the deadhouse has been performed with beneficial outcome.

Noel Paton, Findlay and Watson[80] have recently contributed greatly to our knowledge of the physiologic pathology of tetania thyreopriva. as the above condition is called. The symptoms are not due to any condition affecting the muscles themselves, since they disappear after section of the nerves. Nor are they primarily dependent upon the cerebrum or cerebellum, since ablation of neither abolishes them. This does not imply that secondary involvement of the higher centers never occurs; on the contrary, the epileptiform convulsions and disturbances of equilibrium sometimes observed indicate cerebral or cerebellar involvement, respectively. This leaves some part of the lower neuron reflex arcs as the site of involvement. It is not the afferent neuron, since the tremors and jerkings persist after section of the posterior roots, leaving the afferent neuron as the affected structure.

The foregoing conclusion led Paton and his co-workers to compare the response of muscle and nerve to electric stimulation in normal and parathyroidectomized animals. Although there are considerable variations in the responses of a normal animal, they are very definitely exaggerated in tetany when either the motor neuron or the muscle itself is stimulated, the exaggeration in the latter case being dependent upon alterations in the neural structures (nerve endings) in the muscle. The increased electric excitability can not, however, be taken as a measure of the severity of the condition, for it may be no more marked in cases in which there is involvement of the cerebral hemisphere (causing epileptiform fits) than in milder cases.

As to the cause of the symptoms, many possibilities have to be considered. In the first place, no direct relationship exists between the thyroid and parathyroid in this connection. One cause might be the absence of some substance which checks the activity of the nervous system, some chalone in Schäfer's sense. That such is not the case is shown among other things by the fact that bleeding and transfusing normal saline immediately removes the symptoms for some time. Moreover, the metabolic disturbances go on when the nervous symptoms are slight. It had previously been thought by W. G. Macallum[81] that, since symp-

toms like those of tetany can be induced by deficiency of calcium in the body and the symptoms of parathyroidectomy relieved by administration of this cation, calcium deficiency is the cause of the symptoms. While not denying that these ions may have some relationship to the symptoms, Noel Paton ascribes them to *intoxication by guanidine* (page 605). The evidence is as follows: (1) Guanidine and methyl guanidine administered to normal animals produce symptoms that are identical with those following parathyroidectomy. (2) There is a marked increase in the amount of these substances in the blood and urine of parathyroidectomized dogs and in the urine of children suffering from idiopathic tetany. (3) In certain cases the serum of parathyroidectomized dogs acts upon the muscles of the frog similarly to weak solutions of guanidine and methyl guanidine. (4) There is a striking similarity in the relative amounts of the nitrogenous metabolites in the urine of parathyroidectomized dogs and of normal animals injected with guanidine.

It is concluded that the parathyroids control the metabolism of guanidine "by preventing its development in undue amounts. In this way they probably exercise a regulative action upon the tone of the skeletal muscles." Since it is similar with regard to its characters and metabolism to the condition following thyroidectomy, it is believed that disease of the parathyroids is the cause of idiopathic tetany.

The Relationship of the Parathyroid with Other Endocrine Organs

We know very little of the relationship of the parathyroid with other endocrine organs. Vincent and others have stated that after removal of the thyroid itself enlargement of the parathyroid may occur with the formation of colloid material between the rows of cells, but the conclusion that this represents a vicarious function between the thyroid and parathyroid glands is not generally accepted. The supposed relationships among the parathyroid and the pituitary and adrenal glands are also based upon uncertain evidence.

CHAPTER LXXXIV

THE PITUITARY BODY

Structural Relationships

Situated at the base of the brain and lying in the sella turcica, the pituitary body in man does not weigh much more than half a gram. It is connected with the brain by a funnel-shaped stalk, the infundibulum. On account of a natural cleft, which runs across the gland in an oblique plane, it is an easy matter to split it into two portions, an anterior, or pars glandularis, and a posterior, or pars nervosa. This cleft in the case of man is usually found to be more or less broken up into isolated cysts containing a colloid-like material, and it represents the remains of the original tubular structure from which the pars glandularis is developed; namely, a pouch growing out from the buccal ectoderm.

On histologic examination it will be found that the *pars glandularis* consists of masses of epithelial cells with large sinus-like blood capillaries lying between them. These blood vessels are very numerous, so that in an injected gland this portion of the pituitary stands out very prominently. The vessels are derived from about twenty small arterioles that converge toward the pituitary from the circle of Willis, and enter the gland by the infundibulum or stalk by which the gland is connected with the base of the brain. Three types of cell can be differentiated: nonstaining (chromaphobe) and granular (chromaphil), of which latter there are cells with acid-staining and others with base-staining granules, the former being by far the more numerous (Schäfer).[60] In some animals such as the cat, the cells of the pars anterior are arranged around the blood sinuses in rows as in a columnar epithelium. The cells with acid-staining granules are said to become much increased in number in pregnancy and also in the enlarged gland of acromegaly (see page 772). After thyroidectomy it has been observed that colloid-like masses accumulate in the pars glandularis, the cells sometimes arranging themselves around these masses as in the thyroid gland. The colloid, however, contains no iodine.

The posterior part of the gland, or *pars nervosa,* is composed almost entirely of neuroglia, cells, and fibers, usually with some hyaline or granular material lying between them, particularly in the neighborhood

of the infundibulum, into which it may be traced. It is beli the active principle of the gland is represented by this mater blood supply of the pars nervosa is relatively scanty.

Between the pars nervosa and the intraglandular cleft above to is a layer of cells differing from those of either the anteri posterior lobe. This layer of cells constitutes the so-called *p media.* The cells are somewhat like those of the pars glandular that they are distinctly granular, the granules being of the ne variety, that is to say, they stain with neither basic nor acid dy defined vesicles containing an oxyphile colloid material are oft

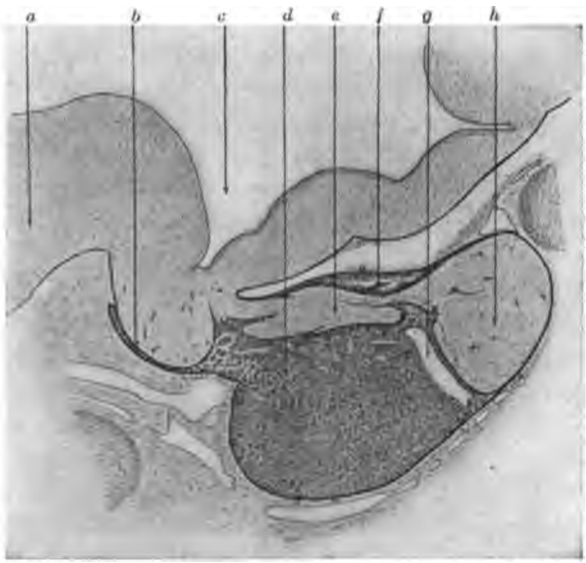


Fig. 196.—Drawing from a photograph of a mesial sagittal section through the pi of a human fetus (5th month): *a*, optic chiasma; *c*, third ventricle; *d*, pars glandula dibulum surrounded by epithelial cells; *f*, pars intermedia; *g*, intraglandular cleft; *h*, (Herring, from Howell's *Physiology*.)

between them. The blood supply is much less abundant than tl pars glandularis. Although well separated by the cleft from glandularis, the pars intermedia is not well separated from nervosa, because many of its cells extend for some distance int ter between the neuroglial fibers. Certain of the cells in the p media may be seen in various stages of conversion into globula bodies, or a granular mass of material may appear in them. case, the cells ultimately break down, setting free the hyaline or material, which is believed to be the origin of the similar ma ready described as existing between the neuroglial fibers of nervosa and therefore ultimately finding its way by the infu

into the third ventricle of the brain. These hyaline globules are greatly increased after thyroidectomy. It should be mentioned, finally, that at the margin of the intraglandular cleft the intermediary and anterior portions of the pituitary come together, although the cells of each can readily be distinguished on account of their staining properties. This pars glandularis et intermedia also extends as a thin layer over part of the pars nervosa and around the neck of the gland at the infundibulum. These relationships are well shown in the accompanying diagram (Fig. 196).

Functions

Concerning the function of the pituitary, it may be said in general that the anterior lobe has an important relationship to the nutritive condition of the body during growth, especially of the skeletal structures, and that the posterior lobe produces a very active autacoid having to do with the physiologic activity of unstriped muscle fiber. The pars intermedia seems to be associated with the posterior lobe in the production of this autacoid. The function of these two parts will therefore be considered together.

Function of the Pars Glandularis.—The facts concerning the function of the pars glandularis have been gleaned largely by observing the effects produced by partial or complete removal of the entire pituitary, justification for ascribing to the removal of the anterior, rather than the posterior, lobe the results that are obtained being furnished by control experiments, in which by removal of the posterior lobe alone similar effects are not observed.

Complete removal of the pituitary is almost invariably fatal, the condition being called *apituitarism.* Two operative procedures have been employed for the removal of the gland. One of these, elaborated by Cushing and his pupils,[82] consists in trephining the skull and elevating the temporal lobe of the cerebrum so as to expose the gland. The other, elaborated by Horsley,[83] consists in approaching the gland through the orbital cavity. Although there is some danger of injury to nervous tissues by the intracranial method, its results are more dependable since the gland is actually exposed to view before being removed.

Most hypophysectomized animals die within two or three days, unless they are very young. This longer survival of young animals is ascribed to the presence of accessory pituitary material situated in the dura mater lining the sella turcica. The most extensive observations have been made on dogs. On the day following the operation the animal appears about normal, but it gradually becomes less active, refusing food and responding slowly to stimulation. It gradually gets weaker and weaker; muscu-

lar tremors may appear, the respiration and pulse become slow, the back arched, the temperature subnormal; and, usually within about forty-eight hours, coma develops and the animal dies in this condition. When the symptoms are less acute and death does not occur so early, it is believed by Cushing either that small portions of the gland have been left behind or that some vicarious activity of other organs has developed to replace that of the pituitary.

When only a part of the pituitary is removed either unintentionally or intentionally, the symptoms are not nearly so acute, and the condition is known as *hypopituitarism.* It is by a study of this condition that most facts concerning the function of the anterior lobe have been learned. When the operation is performed on young animals, they fail to grow properly; the milk teeth and the lanugo are retained; the epiphyses do not ankylose; the thyroid and thymus glands are enlarged; and the cortex of the suprarenal and the sexual organs fails to develop. The animal, though small, becomes very fat and may therefore increase in weight. There is distinct evidence of mental dullness. From these results it is concluded that *the anterior lobe of the pituitary produces autacoids having to do with the development of the skeletal and other structures of the growing animal.* That this autacoid is not derived from the posterior lobe is evidenced by the fact that partial injury of this lobe, or indeed its entire removal, is not followed by similar symptoms.

Closer examination of the metabolic function in hypophysectomized animals has shown that there is a marked depression in the respiratory exchange of oxygen and carbon dioxide, and that the ability to metabolize carbohydrate becomes heightened; that is to say, the animal without developing glycosuria can tolerate a larger quantity of sugar than the normal animal. This effect on carbohydrate metabolism may however be associated not so much with the function of the anterior lobe as with that of the posterior, for, as we shall see later, Cushing and his pupils have found that extract of the posterior lobe has a marked effect on the assimilation limit of carbohydrate.

Attempts have been made to graft the pituitary, especially the anterior lobe, into various parts of the body. It has been found, however, that within a few days the grafts atrophy and disappear unless there has been complete removal of the pituitary itself, in which case the graft may remain for a month or so and the otherwise fatal outcome of hypophysectomy be warded off. Sometimes, where the graft has remained for a longer time, it is said that a temporary increase in the growth of the animal has been noticed.

Other observers have investigated the effects in normal animals of continuous oral administration of pituitary substance or of subcutaneous

injection of extract. The earlier results were indefinite and confusing, but recently Brailsford Robertson[84] has succeeded in isolating from the anterior lobe a substance called *tethelin,* which accelerates growth in young animals and is thought to have a possible value in hastening the healing process in wounds.

Tethelin is precipitated by dry ether from an alcoholic extract of the carefully isolated anterior lobes. It contains 1.4 per cent of phosphorus and nitrogen in the proportion of four atoms for every atom of phosphorus, two of the nitrogen atoms being present as amino groups and one in an imino group. The effects on growth of mice are in every particular like those of the administration of anterior lobes, and consist in retardation of the first portion of the third growth cycle,* followed by acceleration of the latter portion of this cycle. When fully grown, tethelin-fed mice also differ from normal animals in being smaller in size but of greater weight, with a distinct difference in the condition of the coat. Normal animals at fourteen months of age have "shaggy, staring and discolored coats," whereas in tethelin-fed animals they have the glossy and silky appearance of young animals. During growth, normal animals display a greater variability in weight than tethelin-fed animals.

Extraordinary effects have been observed by Clark[85] to be produced by feeding laying hens with pituitary gland. Thus, by giving to one-year-old hens, in addition to their usual food, 20 milligrams of fresh pituitary substance for four days, it was found that the average daily number of eggs laid by a batch of 655 hens was raised from 273 during the four days preceding the pituitary feeding to 352 during the four days of the administration, these results being obtained at a time of year when the natural egg-production of the hens was diminishing. It was further observed that not only is the output of eggs greatly increased as a result of the pituitary feeding, but likewise their fertility, for in another experiment in which 35 hens were kept along with two cockerels of the same breed, not only was the output of eggs increased (from 18 up to 33), but the fertility of the eggs was greatly enhanced.

Functions of the Posterior Lobe or Pars Nervosa.—As already mentioned, excision of this part of the pituitary can be tolerably well withstood by the animal, so much so indeed that from its behavior after the operation we can conclude little as to the function of the lobe. On the other hand, extracts of the posterior lobe injected into normal animals produce effects that are very striking, indicating that the main function

*Robertson has contributed valuable and very extensive data on the normal curve of growth of white mice kept under carefully controlled conditions. Three growth cycles are present: the first attains its maximum velocity between seven and fourteen days after birth; the second, between twenty-one and twenty-eight days; and the third about six weeks, after which the velocity decreases progressively, until further growth ceases between the fiftieth and sixtieth weeks succeeding birth.

of this lobe is production of an autocoid. The extracts have more or less an epinephrine-like action. Such extracts, rendered protein-free and sterilized, are obtainable on the market under the various names of pituitrin, hypophysin, etc. From them a crystallizable material has been obtained, but this is probably a mixture of various substances. In discussing the functions of these various extracts, it must be remembered that the intermediary part (pars intermedia) is included with the posterior lobe in their preparation.

Although the effect of pituitary extract *on plain muscle fiber* (and on glandular tissue) appears, on first sight, to be very like that produced by epinephrine, it has been found on closer examination that the two substances really act in different ways. The rise in blood pressure produced by pituitary autacoid is likely to be more prolonged than that produced by epinephrine. It stimulates increased cardiac activity, but after the vagi have been cut or sufficient atropine administered to paralyze them, the pituitary autacoid continues to stimulate the strength of the heartbeat without producing the acceleration noted with epinephrine. Whereas epinephrine has little or no action on the coronary vessels or on those of the lungs, pituitary autacoid usually produces constriction of both types of vessel; and on the renal arteries the actions of the two autacoids are entirely different, for epinephrine has a marked constricting effect, while the pituitary autacoid produces dilatation.

Another striking difference in the extracts from the two glands is revealed by repeating the injection after the effect of a previous one has completely passed off. With epinephrine the original effect is reproduced; with pituitrin, on the other hand, the effect of the second injection is very often the reverse of that of the first; that is to say, the blood pressure, instead of rising, may fall, or the rise be very much less marked. Whether this effect of the second dose is caused by the action of an autacoid having a chalonic rather than a hormonic influence, or whether it is due to a reversed effect of the same hormone, it is impossible at present to say. The chalonic effect in any case is much more evanescent than the hormonic, and it is not caused by cholin, as some have suggested. The effect of epinephrine, it will be remembered, is abolished by ergotoxin and apocodeine. These drugs, on the other hand, have no influence on the action of pituitrin. The difference in action between the two autacoids is usually explained by assuming that the epinephrine acts on the receptor substance associated in some way with terminations of the sympathetic nerve fibers in involuntary muscle, whereas *pituitrin acts directly on the involuntary muscle fibers themselves.*

Other types of involuntary fiber are also acted on by pituitrin. The uterine contractions for example are stimulated (Fig. 197); so are those of

the intestine (in contrast to the inhibiting effect of epinephrine), and of the bladder-ureter musculature. Dilatation of the pupil of the excised frog

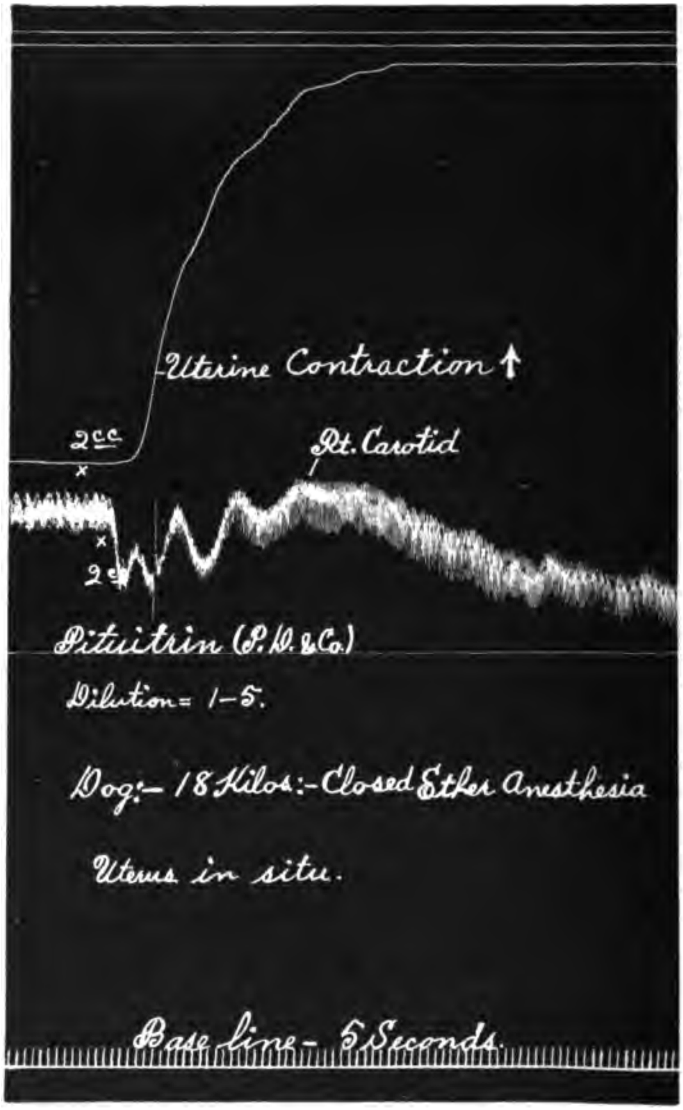


Fig. 197.—Tracing showing the action of pituitrin on the uterine contractions and blood pressure in a dog. Made by Barbour's method. (From Jackson.)

eye is produced. The effect of pituitrin on the muscle of the bronchioles is shown in Fig. 198.

The *glands* on which the pituitrin has the most pronounced action are

the mammary glands and the kidneys. *The effect on the kidney* is evidenced by the remarkable increase in the urinary flow following injection of the pituitrin. This diuresis might of course be due merely to the vasodilatation that we have seen such extracts produce—a vasodilatation which is all the more marked because the vessels elsewhere in the body undergo constriction. But pituitrin continues to cause increased urinary outflow in the absence of any demonstrable vascular change; it also acts after the administration of atropine, so that it is considered by most observers to act on the excretory epithelium of the convoluted tubules

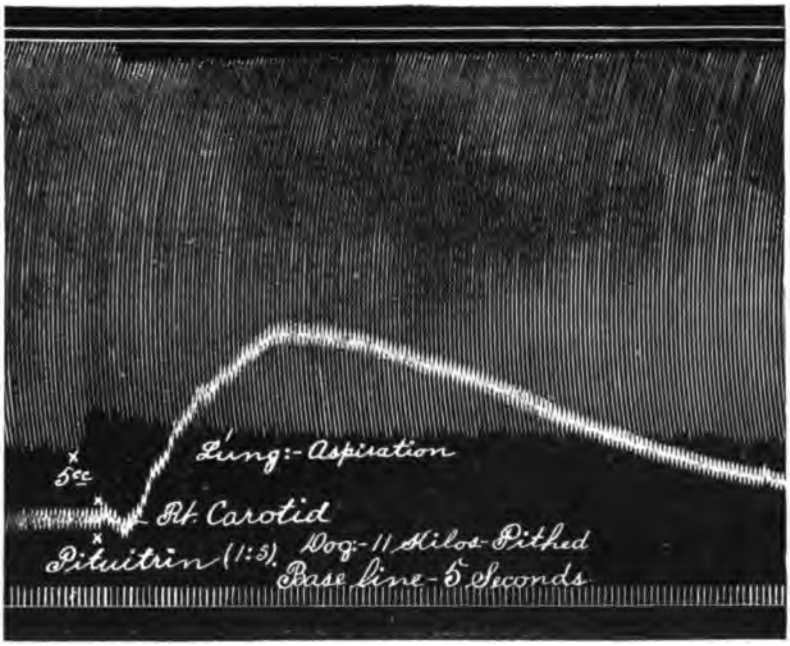


Fig. 198.—Tracing showing the constricting action of pituitrin on the bronchioles and its effect on blood pressure in a spinal dog. (From Jackson.)

in much the same way as certain diuretics, like diuretin. This renal hormonic action of pituitrin would appear to be analogous with that of secretin on the epithelium of the pancreas. Another reason for believing that the secretory hormone is independent of that producing vasodilatation of the renal vessels is the fact that a repeated dose of pituitrin, although, as we have seen, it usually has a depressor action on the blood vessels, still produces a stimulating effect on the excretion of urine. The value of pituitrin as a diuretic in clinical practice is now well recognized.

The effect on *milk secretion* is best demonstrated by placing a cannula

in the mammary ducts so that the milk may freely flow out. By observing the rate of outflow during the injection of pituitrin, it will be found that a remarkable increase occurs. After this increased secretion has ceased, however, the injection of more pituitrin has no further effect, indicating that the influence of the first injection must have been, not so much to stimulate the secretion of milk, as to accelerate the outflow of that which previously had been secreted and had collected in the alveoli and ducts. This effect explains why the pituitary galactagogue should have very little if any effect on the total production of milk or on the total amount of fat and other constituents contained in it. Histologic examination of sections of a resting mammary gland and of the same gland after administration of the pituitrin, bears out the above interpretation of the action. Alveoli in the resting state will be found largely distended with milk and the epithelium flattened against the basal membrane, whereas alveoli from the gland after pituitary activity show small shriveled-up alveoli, containing little milk, and with epithelium that is well marked and stands out prominently from the basal membrane.

These facts taken together indicate that pituitrin stimulates the muscular fibers of the ducts of the mammary glands, thus squeezing out the milk contained in them. Muscular fibers have been described as existing between the basal membrane and epithelial cells, much in the same way as they do in the case of the sweat glands. At least Schäfer has succeeded in demonstrating in this position rod-shaped nuclei which probably belong to muscular fibers.[60] By their contraction, the milk in the alveoli is expelled into the ducts. It has also been found that pituitrin stimulates the secretion of cerebrospinal fluid, and that this stimulation is independent of a rise in blood pressure.

Pituitrin has a distinct *effect on carbohydrate metabolism.* After its intravenous or subcutaneous injection, a marked lowering in the tolerance for sugar is observed (page 652), usually to such an extent that glycosuria becomes established. Cushing and his pupils have concluded that the posterior lobe contributes an autacoid which stimulates the utilization of sugar in the body. Confirmatory evidence for this view is furnished by the observation that mechanical stimulation of the posterior lobe, such as is produced by puncturing it with a needle, is followed by a temporary glycosuria, which is said to be as pronounced as that following puncture of the diabetic center (page 672), provided glycogen is present in the liver. The production of this carbohydrate autacoid would appear to be under the control of the sympathetic nervous system, for it has been found by Cushing and others that stimulation of the superior cervical ganglion, which has been known for many years to be frequently followed by glycosuria, has this effect only provided the posterior

lobe of the pituitary is intact. Even surgical manipulation of the pituitary may excite a hypersecretion of pituitrin, which would account for the glycosuria often observed after experimental excision or partial destruction of the pituitary. A similar irritation may be set up in disease of the gland.

The glycosuria which is usually observed after partial hypophysectomy soon passes off, to be followed by a permanent condition of increased tolerance for sugar, because now less pituitrin is being produced. It is said that during the stage of increased tolerance diabetes can not be produced even by excision of the pancreas. The glycosuria produced by irritation of the posterior lobe is accompanied by a marked polyuria (diabetes insipidus), which may outlast the glycosuria.

A. *B.*

Fig. 199.—*A*, To show the appearance before the onset of acromegalic symptoms; *B*, The appearance after seventeen years of the disease. (After Campbell Geddes.)

Clinical Characteristics

Because of their importance from a physiologic standpoint, we shall now proceed to review briefly some of the more important facts that have so far been brought to light by clinical observations. The pathologic condition most frequently observed affecting the pituitary is an adenomatous growth particularly located in the anterior lobe. Besides producing general symptoms of pressure, such as diminution of the visual field and, perhaps, headache, a shadow can usually be observed when the patient is examined by means of the x-rays. General symptoms, commonly ascribed to a hypersecretion of the autacoid of the anterior lobe of the pituitary—*hyperpituitarism*—begin sooner or later to show them-

selves. These symptoms are almost exactly opposite in character to those observed in animals after removal of this portion of the gland. Thus, the bones of the extremities and of the face become stimulated to increased growth, so that if the patient is young, and the epiphyses therefore not ossified, remarkable elongation of the long bones occurs, producing the condition known as *gigantism*. On the other hand, if the disease does not develop until after ossification is complete, its effects become most marked in the bones of the face, the lower jaw becoming

Fig. 200.—Hand of a person affected with acromegaly.

enormously hypertrophied and the supraorbital ridges very prominent. The long bones also become enlarged at their extremities, and there may be some increase in length of the vertebral column, although the stature does not increase because of kyphosis (curvature of the spine). The condition is called *acromegaly*. Nutritive disturbances of the skin and hairs also become marked, causing the skin to become dry and yellowish, and the hairs to undergo abnormal increase over the body. An early symptom of the condition is a failure of the sexual power (Figs. 199 and 200.)

After a time the disease begins to affect the pars intermedia et nervosa, and disturbances in carbohydrate metabolism come to be observed, consisting usually in a diminished tolerance accompanied by glycosuria, in the early stages of the disease, followed by increased tolerance in the later stages. The glycosuria is usually accompanied by marked polyuria.

It should be observed that sometimes tumor of the pituitary has been found to exist *postmortem* though none of the above symptoms had been recorded during life. In these cases it is probable that the disease from the start had been of such a nature as to produce a tendency to hypopituitarism rather than hyperpituitarism, for the symptoms are very like those observed in animals after partial or complete removal of the gland. If the condition commences before adolescence, the body fails to grow, although the child may continue to increase in weight because of the remarkable deposition of fat in the tissues. Sexual development is strikingly interfered with, and the secondary sexual characteristics fail to show themselves. In boys, for example, the pubic hairs fail to extend up to the umbilicus; and the hairs on the chin do not develop, whereas the hair of the scalp grows profusely. The bones remain of the female type, and a broad pelvis, rounded limbs, small feet and hands are often observed. In these cases there is usually excessive tolerance for carbohydrates, which may explain the adiposity, sugar being converted into fat. In the light of the experimental results, the effect on carbohydrate metabolism may be explained as due to involvement of the posterior lobe. Mental development is retarded, and psychic derangements are sometimes observed.

Where the hypopituitarism does not develop until after adolescence, some of the above symptoms will of course be missed, but many will be observed, such as dryness of the skin, loss of hair, and the tendency in the male to adopt certain of the female characteristics, particularly with regard to the growth of hair. Obesity and increased tolerance for sugar are also evident, and pigmentation of the skin, something like that of Addison's disease, is said often to be a prominent feature. Operative interference in the early stages in many of these cases is of undoubted benefit, as is shown by the brilliant work of Harvey Cushing, to which the reader is referred for further information.

The Relationship of the Pituitary Gland with Other Endocrine Organs

The relationship of the pituitary gland with other endocrine organs seems to be an intimate one.

1. **With the Thyroid and Parathyroid Glands.**—That enlargement of the pituitary occurs after thyroidectomy in man has been known for a

considerable number of years. The enlargement affects more particularly the pars anterior, although changes are also described in the pars intermedia et nervosa. Accompanying the enlargement of the anterior lobe, vesicles containing colloid-like material often become developed in it, but even after the hypertrophy has proceeded to a considerable degree, this colloid does not contain iodine, nor does an extract have the same physiologic effect as one of the thyroid gland. It can not replace thyroid extract in the treatment of patients with goiter or myxedema, or ameliorate the symptoms produced in animals by the removal of the thyroid gland. Deposition of colloid-like material in the pars anterior also occurs in myxedema. Histologic changes in the pars intermedia et nervosa, although less pronounced than in the pars anterior, are nevertheless said to be perfectly distinct following thyroidectomy, and to consist in an increase in the hyaline and granular masses which have already been described as present to a certain extent in the normal gland.

Less direct evidence of an association in function between the pituitary and the thyroid is furnished by the similarity of the effects produced on the sexual functions and on the general development of young animals by the removal of either gland. In both cases the animals fail to grow properly; the sexual organs remain undeveloped; and the mental functions are infantile in type. In hypophysial deficiency, however, extreme adiposity is likely to be more marked than is the case in cretinism.

2. **With the Sexual Organs.**—That the pituitary gland has much to do with the development of the sexual organs has already been shown. Further evidence of a relationship between the sexual glands and the pituitary is furnished by the following observations. After castration enlargement occurs in the pituitary, and on histologic examination the gland is found to contain a large number of oxyphile cells, particularly in the pars anterior. This influence of the sexual glands on the pituitary is believed to depend on the interstitial cells present in them, for it has been found that if the ovary or testis is transplanted into other parts of the body after the castration, the changes in the pituitary do not occur, although, as we shall see, the transplanted gland becomes entirely atrophied except for the interstitial cells. The enlargement of the pituitary during pregnancy—an enlargement which often brings it to two or three times its normal weight—is further evidence of its association with the ovary.

3. **With the Suprarenals.**—Association of function is suggested in this case by the fact that extracts of suprarenal and pituitary have very much the same effects on involuntary muscular fiber and glandular structures, and it is said that the two extracts mutually facilitate each other's action in this regard. It should be remembered, however, that pituitrin

and epinephrine do not appear to act on exactly the same peripheral mechanism (see page 767).

4. **With the Isles of Langerhans.**—Since pituitrin affects carbohydrate metabolism, which is thought to be primarily controlled by the Isles of Langerhans, it is claimed by some observers that a relationship also exists between the pituitary and these structures. Injections of duodenal extracts are also said to cause a hypersecretion of pituitrin into the cerebrospinal fluid.

CHAPTER LXXXV

THE PINEAL GLAND AND THE GONADS

THE PINEAL GLAND

This peculiar structure lies between the anterior corpora quadrigemina, and weighs about two-tenths of a gram. It is largest in the early years of life, and undergoes retrogressive changes after puberty. Microscopically it consists of epithelial cells arranged loosely in trabeculæ, with large sinus-like capillaries between them; neuroglia and sometimes muscle-fiber cells are also present. Curious globules of calcareous matter (brain-sand) are also found, especially in the pineal gland of man. The gland is developed from an evagination of the third ventricle, and it is homologous with the so-called median eye of reptiles.

The *functions* of the pineal gland are obscure. In cases where its extirpation has been successfully accomplished (in the fowl), it has been found that the body growth is stimulated and the sexual characteristics developed more quickly. This result would seem to indicate that the clinical observation that tumors of the pineal gland are associated in young boys with abnormal growth of the skeleton and with the early development of the secondary sexual characteristics, depends on the fact that a condition of hypopinealism is produced by the growth of a tumor. The immediate effects of the injection of extract of pineal gland are not characteristic, consisting merely of a fall in blood pressure, which is, however, obtainable when an extract of practically any cellular organ is injected. Prolonged administration of an extract to growing animals is said to accelerate the growth and to bring about a precocious development of the sexual organs; but this result is somewhat difficult to interpret, for, as we have just seen, similar changes occur after experimental removal of the gland.

THE GONADS OR THE GENERATIVE ORGANS

The Generative Glands of the Male

The structures which are responsible for the well-known influence of the testicles on the development of the male sexual characteristics are the so-called interstitial cells of Leydig, which consist of polygonal-shaped epithelial-like cells, with well-marked nuclei and nucleoli. Lipoid

granules, staining black with osmic acid, are also present in the cytoplasm. The degree of development of the interstitial cells varies in different animals, being marked in the cat and man and ill-marked in the rat and rabbit. In animals which show seasonal changes in sexual activity, the cells are most prominent between the periods of sexual activity, when the semeniferous epithelium is less evident. They also become prominent in cases where the semeniferous epithelium is atrophied, either as a result of disease or following ligation of the vas deferens done in such a way that the artery and nerves to the testicles are not included in the ligature. When the testicle or a portion of it is grafted into another part of the body, the semeniferous epithelium degenerates, but the interstitial cells remain alive and become quite prominent. It is believed that the interstitial cells are responsible for the production of an autacoid that has to do with the development of accessory sexual characteristics.

The effects of castration are not significant in animals below the vertebrata. In all of these, however, they are very pronounced. The castrated male frog fails to show development of the thumb pad, but this development immediately ensues if portions of testis from another frog be placed in the dorsal lymph sac. In birds the results are more pronounced; in the castrated male chick the comb, spurs, wattles, etc., fail to develop, but will usually do so if some testis from another bird is transplanted into its tissues. In mammals the effects are most striking in animals that develop marked male characteristics, such as the growth of antlers in stags. These fail to develop properly and are prematurely shed after castration. In man also, as is well-known from a study of eunuchs, castration has a very profound effect. Hair fails to grow on the face; the larynx remains undeveloped; the epiphyses are a long time in ossifying, so that the stature may become great, but at the same time the limb bones may be more delicate than usual; the sutures of the skull are slow in closing; and the whole architecture of a castrated male comes to be very like that of the female. Confirmatory evidence of the influence of the testicles on the development of secondary sexual characteristics is afforded by the observation that malignant tumors of the testes in boys are associated with the premature development of the secondary sexual characteristics, and that these may recede after the removal of the tumor.

As a result of castration, interesting changes have also been observed in other ductless glands. Thus, the suprarenal cortex and the thymus become enlarged, whereas the thyroid and pituitary become atrophied. The metabolic functions also become tardy, as is evidenced by a tendency to the deposition of fat.

When the castration is performed on an adult man, the above changes in the sexual characteristics are of course not so evident, although the prostate, etc., atrophy. The effect on the metabolic functions is, however, very marked, there being a striking tendency to increased formation of fat. It is interesting that accompanying this there should usually occur a lowering of the assimilation limit for carbohydrate, so that glycosuria is very readily induced. We can not assume, therefore, as Cushing has done in the case of hypopituitarism, that the fat deposition is attendant upon an improper combustion of carbohydrate.

These remarkable effects of castration have naturally prompted observers to study the influence of injection of testicular extract on the development of sexual characteristics in different animals, but the results have in general been considered to be negative in character.

The Female Generative Organs

It is well known that, besides their function in producing ova, the ovaries also produce autacoids that have to do not only with the fixation of the embryo *in utero,* but also with the changes that occur during pregnancy in the maternal organism. It is however at present uncertain as to where these autacoids are produced in the ovary. The two most likely sources are the stroma cells and the corpus luteum. In the stroma of the ovary of certain animals, groups of cells have been described having a different appearance from those of ordinary stroma cells. They have been called *the interstitial cells* of the ovary, and are believed to be analogous with the similar structures found in the testicle. It is possible, however, that these interstitial cells are nothing more than cells derived from previous corpora lutea. The latter are formed by proliferation of the follicular epithelium which remains after extrusion of the ovum, and by the ingrowing into the follicle of the so-called theca cells and blood vessels. The fully developed corpus luteum in most animals consists of cells arranged in trabeculæ converging toward the scar which formed at the place where the follicle had burst. The luteal cells, as they are called, are characterized by containing considerable quantities of lipoid material.

That the ovary produces some autacoid is evidenced by both clinical and experimental observations. Thus, if both ovaries are removed in a young animal (oöphorectomy or spaying), it is well known that not only does the uterus fail to develop properly, but the external changes characteristic of puberty in the female fail to materialize, although actually the general effects are not so pronounced as they are in the male after castration. Menstruation does not set in; the mammary glands fail to develop; and there is a tendency for the hair to grow as in the male.

When the operation is performed in adult life, the changes are not very pronounced, except that menstruation ceases and the uterus and mammary glands atrophy. Metabolism also becomes altered, causing a tendency to the deposition of fat, and in the case of the human animal at least, there is frequently evidence of mental disturbance.

Attempts to acquire more definite information regarding the physiologic effects of the ovarian autacoid have recently been made by Schäfer and Itagaki.[60] Extracts were prepared from the corpus luteum or Graafian follicles or from the hilum ovariae, and observations were made on the effect produced on the behavior of the chief forms of unstriated muscle by adding the extracts to isolated preparations of uterus or intestine or by injecting the extracts into animals. Applied to the isolated preparations, extract of follicular tissue or of liquor folliculi was found to increase the force and rate of the rhythmic contractions of the uterus as well as its tone, whereas inhibition was produced when extract of the hilum was used. Extract of corpus luteum, when injected into the veins, was found to cause the uterus to increase its contraction or if quiescent to begin contracting. It was further noted that extracts of hilum caused a fall in arterial blood pressure, whereas those of corpus luteum had little or no effect. It would appear from these observations that the extracts contain two different autacoids, one having a hormonic and the other a chalonic action on plain muscular fiber.

Extract of corpus luteum when intravenously injected also stimulates the outpouring of the milk from the mammary glands, although not so markedly so as extract of pituitary gland. This pituitary-like action is not obtained with extracts of ovary that do not contain corpora lutea. Besides being concerned in the outpouring of milk, corpus luteum has also been shown to be related in some way to the development of the mammary gland during pregnancy. These glands become developed in young virgin rabbits after the continuous administration for a month or so of extract of corpus luteum, and they also develop in unimpregnated animals when the corpus luteum is made to develop by artificial means such as puncturing the Graafian follicle. Furthermore, destruction of the corpora lutea in a pregnant rabbit arrests development of the mammary glands. The corpus luteum has also an important function in connection with the formation of the uterine decidua and the fixation of the embryo. Thus, after destruction of the corpus luteum at an early period in pregnancy, the embryo fails to become adherent to the uterus.

DUCTLESS GLANDS REFERENCES*

(Monographs)

[58]Vincent, Swale: Internal Secretions and the Ductless Glands, Ed. Arnold, London.
[59]Biedl: The Internal Secretory Organs, Wm. Wood & Co., 1913.
[60]Schäfer, Sir E. A.: The Endocrine Organs, Longmans, Green & Co., New York and London, 1916.

(Original Papers)

[61]Fulk, M. E., and Macleod, J. J. R.: Am. Jour. Physiol., 1916, xl, 21.
[62]Folin, O., Cannon, W. B., and Denis, W.: Jour. Biol. Chem., 1913, xiii, 447.
[63]Cannon, W. B., and Gray, H.: Am. Jour. Physiol., 1914, xxxiv, 232; also with Mendenhall, W. L.: Ibid., 243 and 251.
[64]Hartman, T. H., and others: Am. Jour. Physiol., 1915, xxxviii, 433; ibid., 1917, xliii, 311; ibid., xliv, 353; ibid., 1918, xlv.
[65]Hoskins, R. G.: Am. Jour. Physiol., 1912, xxix, 363; Jour. Pharm. and Exp. Therap., 1911, iii, 93; Am. Jour. Physiol., 1915, xxxvii, 471; ibid., 1916, xli, 513.
[66]Stewart, G. N., and Rogoff, J. M.: Jour. Lab. and Clin. Med., 1918, iii, 209. See full bibliography by Rogoff in this paper.
[67]Elliott, T. R.: Jour. Physiol., 1912, xliv, 374.
[68]Stewart, G. N.: Jour. Exp. Med., 1911, xiv, 377; ibid., 1912, xv, 547; ibid., xvi, 502.
[69]Stewart, G. N., Rogoff, J. M., and Gibson: Jour. Pharm. and Exper. Therap., 1916, viii, 205.
[70]Meltzer, S. J.: Deutsch. med. Wchnschr., 1909, xiii.
[71]Stewart, G. N.: Jour. Exper. Med., 1912, xv, 547.
[72]Cannon, W. B., et al.: Am. Jour. Physiol., 1911, xxviii, 64; ibid., 1914, xxxiii, 356; also Bodily Changes in Hunger, Fear, and Rage, Appleton, 1915.
[73]Cannon, W. B., and Cattell, McKeen: Am. Jour. Physiol., 1916, xli, 74.
[74]Macleod, J. J. R., and Pearce, R. G.: Am. Jour. Physiol., 1912, xxix, 419.
[75]Marine, D.: Personal communication.
[76]Marine, D.: Jour. Exper. Med., 1914, xix, 89.
[77]Marine, D., and Lenhart, C. H.: Jour. Exper. Med., 1910, xii, 311; ibid., 1911, xiii, 455; also Bull. Johns Hopkins Hosp., 1910, xxi, 95.
[78]Marine, D., and Kimball, O. P.: Jour. Lab. and Clin. Med., 1917, iii, 41.
[79]Kendall, E. C.: Boston Med. and Surg. Jour., 1916, 175, 557; also Proc. Am. Physiol. Soc., Am. Jour. Physiol., 1918, xliv.
[80]Paton, Noel and Finlay: Quart. Jour. Exp. Physiol., 1917, x, 203. Paton, Noel, Finlay and Watson, A.: Ibid., 233, 243, 315, and 377.
[81]MacCallum, W. G., etc.: Jour. Exper. Med., 1909, xi, 118; ibid., 1913, xviii, 646; Jour. Pharm. and Exper. Therap., 1911, ii, 421.
[82]Cushing, Harvey: The Pituitary Body and Its Disorders, J. B. Lippincott Co., 1912.
[83]Horsley, V.: Brit. Med. Jour., 1885, i, 111.
[84]Robertson, Brailsford, and Ray, L. A.: Jour. Biol. Chem., 1916, xxiv, 347, 363, 385, 397, 409.
[85]Clark, L. N.: Jour. Biol. Chem., 1915, xxii, 485.

*The numbering is in continuation with that for metabolism.

PART IX

THE CENTRAL NERVOUS SYSTEM

CHAPTER LXXXVI

THE EVOLUTION OF THE NERVOUS SYSTEM

The nervous system of the higher animals consists of the nerve centers, and the nerves with their various interconnecting tracts. The nerve tract and centers are located mainly in the spinal cord and brain, where, by their interlacement, they form an extremely complex structure. The exact position of the centers and the course and connections of the tracts with the centers are problems which, under the title of neurology, have during recent years been contributed to more particularly by the anatomist and the pathologist. The information thus gathered tells us the possible tract or tracts of nerve fibers through which the various centers may communicate either with one another or with the structures outside the central nervous system upon which they act. Since each of these centers may, however, be played upon by influences coming from different regions of the body, it is evident that there must remain, as an equally important aspect of the subject, the investigation of the means by which the various available centers and tracts are brought into communication and action at the proper time. In other words, we must investigate *the functional uses of the available paths.*

We may compare the central nervous system with a telephone system, the exchanges representing the nerve centers, and the wires the nerve trunks. Any incoming wire may be connected by the operator with any outgoing wire, but a knowledge of how each wire runs does not tell us under what conditions the various wires will be connected for transmission of messages. It is the same with the nervous system; the neurologist can tell us how the tracts and centers run, but not the conditions under which they may act together. This it is the duty of the physiologist to ascertain.

Since it is the degree of development of the central nervous system which determines an animal's position in the evolutionary scale, much information concerning the relative importance of the various parts of

it can be gleaned from a survey of the conditions under which the nervous system makes its appearance in the lowest forms of animal life. In the case of unicellular organisms, such as the *ameba*, the application of a stimulus to the surface causes a movement, because the protoplasm of the organism possesses, among its other properties, those of excitability, conductivity and contractility. In the case of multicellular organisms, on the other hand, some cells are set aside and specialized for the assimilation of food, others for movement, others to receive stimuli from the outside, and yet others to compose the tougher tissues which protect the surface of the animal from injury. This location of specific function in specialized groups of cells makes it necessary, for the welfare of the organism as a whole, that some means of communication should be provided between the distant parts of the animal, for otherwise the cells which are occupied in absorbing food would be unable to move away or be protected from harm when some destructive agency approached them, and indeed the moving (muscle) cells could never know when the welfare of the organism as a whole demanded that they should become active.

It is probable that, in some of the lower organisms, the messages transmitted from one group of cells to the others are carried by chemical substances present in the circulating fluid—hormones, as they are called (page 729). For the quick adaptation that is necessary in the struggle for existence, however, such hormones are usually too slow in bringing about the response, and very early in the evolutionary scale we find that certain cells become differentiated for this special purpose. The cells thus specialized constitute the nervous system, their differentiation, as would be expected, being, however, antedated by that of the cells that form the muscular tissues. In the *sponges*, for example, muscle cells become developed from ameboid epithelium and from a layer underneath the external epithelium. These muscle cells contract slowly so as to cause opening and closing of the small mouths, or oscula, on the surface of the sponge in response to movements in the sea water. They are independent of any nervous structures.

In certain *Cœlenterates* the muscle cells respond a little more quickly than in the sponges, and this greater efficiency is found to be dependent upon the appearance of a localized, very primitive nervous system[1]. This nervous system consists of specially modified epithelial cells, or *receptors*, sending branches from their inner ends, which either come in contact with the muscle cells, or *effectors*, or become interlaced so as to form a network. In the region between the receptors and the effectors the network at first serves merely as a structure whereby the entire musculature of the animal can be brought into harmonious action from a single

point on the surface, as, for example, in the case of the sea anemone (No. 2 of Fig. 201). In the jellyfish, which in contrast to the sea anemone is a free moving animal, we find that the receptors are more highly specialize and, therefore, much more sensitive, and that the impulses which they receive are transmitted to a more definite nerve network, capable not only of conveying the excitatory process from one part of the animal to another,

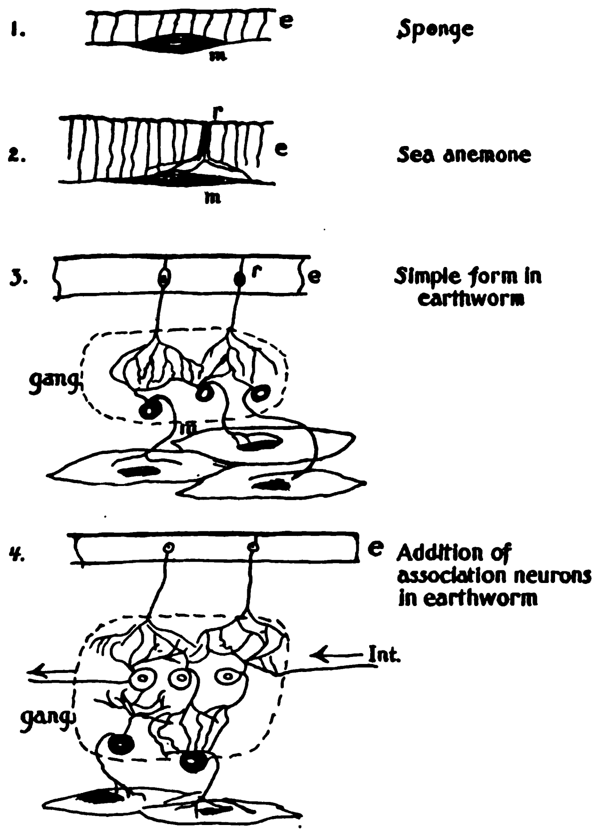


Fig. 201.—Diagram to show gradual evolution of nervous system from an epithelial cell (*e*) and muscle fiber (*m*) in the sponge (*1*) to a specialized epithelial cell or receptor (*r*) and muscle cell in the sea anemone (*2*); then to a receptor and motor neuron joining in a ganglion (*Gang.*), in simple form seen in the earthworm (*3*). Most of the ganglia in this and other segmented invertebrates show also the internuncial or association neurons as indicated in 4.

but also of imprinting on the impulse a characteristic rhythmic activity which brings about the contraction of the bell and the swimming movement of the animal. The network now assumes the function of an *adjuster* as well as a transmitter of impulses.

So far the adjuster is an extremely simple structure, and it is possible that the effector and receptor organs are directly connected by fibers running through it. When we come to the *segmented invertebrates*

(such as the earthworm, crayfish, lobster, etc.,) much more definite specialization of the adjuster occurs, for now this intermediate nervous tissue becomes collected into so-called *ganglia,* a pair existing for each segment and the various pairs being connected by definite nerve structures, constituting the ganglion chain. It is in this group of animals that we have, for the first time, definite evidence of the existence of the *neuron,* which may be considered as the elementary unit of which the nervous system of all the higher animals is built. A neuron may be either sensory or motor, and in both cases it consists of a cell with a nucleus, one long process, called the *axon,* and several short branching processes, called the *dendrites.* The axon in its course may give off a branch, or more, at right angles,—these are sometimes called *collaterals,*—and at its end it may break up into very fine branches called a *synapsis.* In a sensory neuron the impulse is transmitted from the end of the axon to the nerve cell, whereas in a motor neuron it is transmitted in the opposite direction from the cell to the end of the axon (Fig. 203).

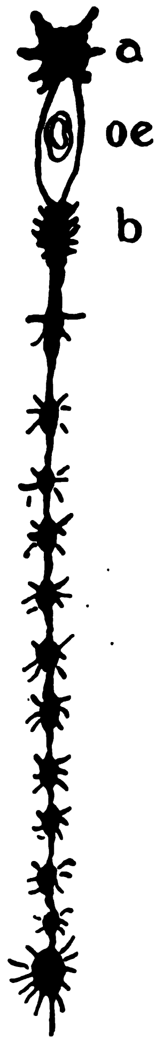


Fig. 202.—Diagram of nervous system of segmented invertebrate; *a,* supra-esophageal ganglion; *b,* subesophageal ganglion; *oe,* esophagus or gullet.

The simplest arrangement of sensory and motor neurons to constitute the nervous system is seen in the earthworm, in which it forms the simplest type of *reflex arc* (Fig. 201, No. 3). The sensory neuron has its cell body in the skin, and its axon proceeds to one of the segmental ganglia, in which are large nerve cells whose thick axons pass out from the ganglion as motor fibers to the muscles of the body wall. The dendrites of the motor neuron and the branching of the termination of the sensory neuron cause a very fine interlacement of nerve fibers in the ganglia, forming a network known as the *neuropile.* The sensory impulse, on reaching the ganglion, is transmitted by the synapsis to the dendrites, probably without the fibers actually joining together; that is, the nerve impulses pass from the one to the other set of branches by contact rather than by transmission through continuous tissue.

By such an arrangement it is evident that the nervous apparatus in each segment could cause a contraction of the muscles of its own neighborhood, but that a stimulus applied to one receptor would be incapable of calling forth a contraction of the muscles of a far distant segment, much less a coordinated contraction of the musculature

of the whole animal such as would be required for locomotion. To render this possible it is necessary that some means of communication become established between the different segmental ganglia. This is effected by ***association neurons,*** each of which, as the name implies, consists of a nerve cell with its dendrites located in one ganglion and of an axon, which passes to the next or even to some more distant ganglion, where it ends by synapsis. The important point to note is that these association neurons do not leave the central nervous system; they merely connect various ganglia.

So far the ganglia of each segment are of equal importance, but if we examine further we shall find that at the head end of the animal several of the ganglia become fused together to form a larger ganglion,

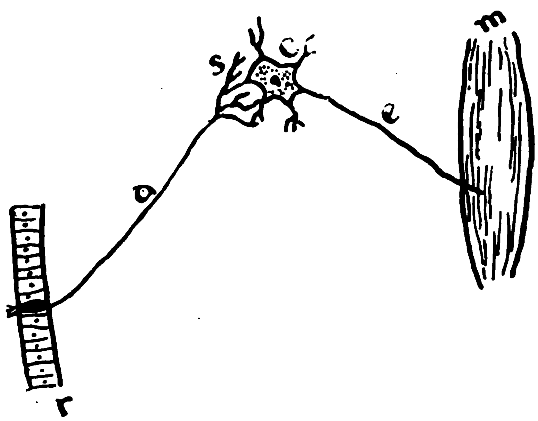


Fig. 203.—Schema of simple reflex arc; *r*, receptor in an epithelial membrane; *a*, afferent fiber; *s*, synapsis; *c*, nerve cell of center; *e*, efferent fiber; *m*, effector organ.

which lies just behind the gullet, and from which fibers proceed around the gullet to unite in front of it in another large ganglion, which usually shows three lobes. These larger ganglia receive afferent nerve fibers from the closely adjacent primitive sense organs for sight, sound and smell, from structures, that is, that are really highly specialized receptors. The cells of the retina and ear have been made capable of reacting to impulses of light or sound instead of those of pain, touch or temperature, to which the receptors of the integument are especially sensitized. They are *distance receptors* (projicient receptors), and it is evident that the nerve reflexes with which they are concerned are of a higher order than those located in the segmental ganglia themselves.

Some of the neurons of the head ganglia are merely motor and act on the muscles of the head end of the animal, but others are purely associa-

tion neurons and proceed down the ganglion chain to terminate by synapses in one or other of the segmental ganglia. These association neurons exercise a dominating influence over the activities of the segmental ganglia, so that they may determine the response of the animal when its safety is threatened by some approaching enemy. When, for example, the stimulus produced by some sight or sound of an approaching enemy is received by the head ganglia, these will transmit impulses down the ganglion chain which so influence the various nerve cells of this chain as to produce in all of them a coordinated action for the purpose of removing the animal from danger. Even should some local stimulant be acting on one or more of the segments, the response may be inhibited on account of stimuli meanwhile transmitted by way of association neurons from the large head ganglia; in other words, the part controlled by the segmental ganglia becomes subservient to the whole through the dominating control of the head ganglia.

This illustrates the beginnings of the *integration of the nervous system;* and as we pass to the study of the higher animals, we shall see that this integration becomes more and more complicated, and that, as it does so, the nerve centers acquire the power of storing away the impressions they receive, which they may afterwards apply to regulate the reflex response. Thus *memory* and *volition* come to find their place in the nervous integration of the animal. The afferent stimulus arriving, let us suppose, at nerve cells controlling the movement of the leg, may fail to cause a response of the corresponding muscles because of impulses meanwhile transmitted by association neurons from higher memory centers, for the animal may have learned by experience that such a movement as the local stimulus would in itself call forth is opposed to its own best interests. This experience will have been stored away in memory nerve centers, so that, whenever the local stimulus is repeated, impulses are discharged from the memory centers to the local nerve centers, and the reflex response does not occur, or is much modified in nature. For storing away these memories and for related psychologic processes of volition, etc., the anterior portions of the nervous system in higher animals become very highly developed so as to constitute ***the brain,*** and the simple chain of ganglia of the invertebrates is replaced by ***the spinal cord.***

As we ascend the scale of the vertebrates, the brain becomes more and more developed, until in the higher mammalia, such as man, very few reflex actions can occur independently of the higher centers which are located in it. The reflex arc now involves, not one nerve center, but several, and of these the most important are located in the brain.

There is thus no essential difference in the general nature of integra-

tion in the nervous system of the lower as compared with the higher animals, but there is a very distinct morphologic difference: in the lower or invertebrate animals the ganglion nerve chain is ventral to the alimentary canal, whereas in the higher or vertebrate, the spinal cord, which takes the place of the ganglia, is dorsal to the alimentary canal. In both groups the head ganglia are dorsal to the alimentary canal, but in the vertebrates these become much more definite in structure, and constitute the brain. This morphologic difference between vertebrates and invertebrates is probably not so fundamental as at first sight it may appear to be, for, as Gaskell has shown, it is possible that the alimentary canal of the invertebrates is really homologous with the central canal of the spinal cord and the ventricles of the brain of the vertebrates. According to this observer, what has really happened in the latter group of animals is that the ganglia have grown up so as to surround the alimentary canal and so constitute a continuous structure, a new alimentary canal being meanwhile provided by the enclosure of a space as a result of ventral downgrowth of the body walls. Although this view has not been generally accepted by biologists, there is no inherent reason why it should not be accepted. It is no more to be wondered at than the well-known fact that a new respiratory system becomes developed in the passage from aquatic to land amphibians.

The fibers of the sensory neurons in vertebrates are collected together to form the *posterior roots* of the spinal cord, and the cell bodies of these neurons are located not on the surface, as in invertebrates, but in the posterior root ganglia, the cells being connected to the fibers by T-shaped junctions. The olfactory nerve is the only one in the higher vertebrates which retains its primitive condition.

In the vertebrate animals the spinal member in the integration of the central nervous system is the motor neuron, the fibers being collected in the *anterior roots*. Toward the cell of this neuron impulses are transmitted, not only from the segment in which it is itself located, but by way of association neurons from other segments or from far distant parts of the central nervous system. In other words, this motor neuron may transmit impulses which cause the muscles to perform local independent movements, which are coordinated with those of adjacent segments and which may be of widely varying types. The motor neuron has therefore very appropriately been called the *final common path,* and it will be one of our main objects later to show the conditions under which several different competing influences may obtain possession of this path.

CHAPTER LXXXVII

THE PROPERTIES OF EACH PART OF THE REFLEX ARC

Having briefly traced the physiologic development of the nervous system, we are prepared to consider in greater detail the peculiar function of each of the parts which enter into the formation of the reflex arc.

THE RECEPTOR

With the advance in animal organization is associated the development of the ability to appreciate and discriminate between external phenomena, special organs called receptors being evolved to receive the stimuli which these occasion. Those receptors which are distributed over the skin of the animal are called *external or exteroceptors,* and are especially adapted to react to such stimuli as temperature, pressure, and pain, but at the fore end of the animal certain receptors become more highly specialized so as to react to stimuli coming from a distance—that is, to stimuli that are not produced by contact of external objects with the surface of the animal. These specialized receptors—sometimes called *projicient*—include the eye, the ear, and the olfactory epithelium. Receptors are also provided in the interior of the organism for the purpose of receiving stimuli dependent upon the activities of the organism itself. They may be called *internal receptors,* and we may further distinguish two groups of them—namely, those which come from the surfaces of the mucous membranes and those which come from the substance of the various organs and tissues themselves, as, for example, from the substance of muscle or tendon.

A receptor may be defined in a general way as a mechanism in which some particular kind of stimulus produces changes that result in the excitation of the nerve fiber with which the receptor is connected, although the stimulus in itself is incapable of exciting the nerve fiber. In other words, as Sherrington puts it, the receptor has the threshold of its excitability raised to every kind of stimulus save one, toward which it is lowered. A nerve fiber, for instance, responds to every kind of stimulus approximately equally; a receptor will also respond to these same stimuli, but with great inequality, since each receptor is specialized to react to one kind of stimulus and to others only when these are very strong.

It is often a difficult matter to determine just exactly what it is in the nature of the stimulus that makes it capable of affecting one receptor and not another; for example, it is often merely a question of the rate of vibration of the stimulus. Light and heat rays are both due to vibration of the ether which fills space. When these vibrations are slow, they stimulate receptors that have been specialized for appreciation of temperature, but when they are rapid and exist as rays of light, they no longer affect the temperature receptors but only the highly specialized receptors of the retina. Similar vibrations of the air in place of the ether cause sound and stimulate the auditory receptors. It is quite likely that the receptors in different groups of animals are attuned to react to different rates of vibration. For example, a cat can hear higher pitched notes than man, and it is possible that the retinas of some animals respond to rays vibrating with a different frequency from those to which the retina of man is adapted. In this connection it is of interest to note that the touch receptors of the skin respond so promptly to stimulation that one hundred vibrations of a tuning fork per second can be felt as separate stimuli, whereas to the ear at this frequency the fork emits a continuous note. The receptors of touch are therefore more prompt in their response than the receptors of the auditory nerve.

When once the receptor has been stimulated, the impulse passes and is transmitted to the nerve centers, where it is translated into a particular sensation. The conditions are really not unlike those which obtain in the case of the various physical instruments used to receive and convert into the electric current stimuli of heat, light, chemical energy, etc. The receiver required to bring about this transformation must be especially constructed in each case, that for light being the actinometer, that for motion the dynamo, that for heat the thermopile, and that for chemical energy the concentration cell. Each of these physical instruments may be considered as a specialized receptor for the purpose of producing an electric current out of other forms of energy.

In accepting the above analogy we must not fail to bear in mind that very feeble stimuli are often able to set in operation nerve impulses that are as potent as those produced by much stronger stimuli. Here again, we have a physical analogue in the case of relay currents, in which a feeble electric current may operate to complete the circuit from independent sources of electric discharge and thus set in motion a much larger amount of energy.

These general considerations of the nature of a receptor naturally lead us to the law of *the specific properties of nerve,* which is to the effect that, however excited, each nerve of special sense gives rise to its own peculiar sensation. Thus, in whatever way the chorda tympani

nerve is stimulated (chemically, mechanically or electrically) during its passage across the tympanum, the sensation evoked is that of taste. And so with the receptor; whatever the means by which it is excited, whether by the particular kind of stimulus for which it is adapted or by excessive intensities of other stimuli, excitation always evokes the same sensation. If the optic nerve or retina is mechanically stimulated, as by pressure against the outer canthus of the eye or by an electric current, the sensation is that of light. Applying these facts to less well-known receptors, such as those of heat and cold, it is interesting to note that stimulation of a "cold spot" by extreme heat or by mechanical or electrical stimuli brings out the sensation of cold.

Properties of Epicritic and Protopathic Receptors

A valuable *grouping of receptors of the skin* has been demonstrated by Head and his pupils by experiments on himself. Head found after section of the skin nerves—of the radial nerve, for example—that deep pressure and pain were still present in the area supplied by the nerve, indicating that these deep sensations are carried by the sensory fibers present in the muscular nerves. In such a paralyzed sensory region the power of general localization is fairly good, although light, touch, temperature and superficial pain are entirely absent in the overlying skin.

In the case of the fingers the nerves of deep sensibility run in the tendons of the finger muscles, so that after severance of the cutaneous nerves and tendons of the hand, all sensibility is gone.

During the regeneration of the cut nerve the cutaneous sensations reappear at two periods: one group, called the *protopathic,* begins to appear in from seven to twenty-six weeks, whereas the other, called *epicritic,* does not fully appear for one or two years.[2] The protophatic sensations are of a distinctly lower order than the epicritic. When they alone are present, there is the sensation of pain, but not that of fine touch; temperature sensations are felt when extreme degrees of heat or cold—above 38° C. or below 20° C.—are applied to the skin, but not for slight degrees; the power of discriminating between two points is almost entirely absent; and the sense of localization is very imperfect. For example, the person will often refer the point that has actually been stimulated to a neighboring normal portion of skin. Protopathic sensibility is more or less distributed in spots, and it is strongly "affective" in character, causing an intense subjective sensation. A stimulus that causes only moderate pain under normal conditions produces in a "protopathic area" a pain that may be intense.

The epicritic sensation, as will be inferred from the foregoing, responds

to finer grades of stimulation. By it we can feel the lightest touch and can discriminate the finest grades of temperature between 26° and 37° C. The power of localization of the stimulus and the ability to discriminate between two points also return with epicritic regeneration.

In the spinal cord the nerve fibers carrying one *kind* of sensation are grouped together, in the sense that pain sensations, whether deep or protopathic, run in the same column in the cord. Likewise temperature sensations, whether protopathic or epicritic, run together.

The Peculiarities of Each of the Separate Sensations

Temperature.—The receptors for temperature are arranged in groups, some being sensitized for heat, others for cold. These groups of receptors are called *heat and cold spots.* They can be very easily detected on an

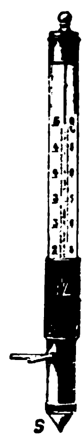
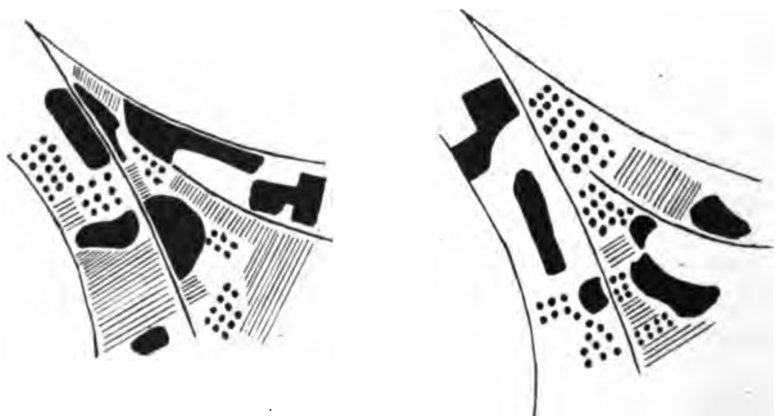


Fig. 204.—Thermoesthesiometer.

area of skin by means of a pointed hollow vessel, through which water is made to flow at a temperature a little below or a little above that of the skin. The instrument is called a thermo-esthesiometer. On a part of the skin where there are no heat and cold spots, the thermo-esthesiometer will elicit no sensation either of heat or of cold. This is charted on an outline drawing of the part as a neutral spot. At other places it will call forth a sensation of heat, indicating the presence of heat spots, or at others a sensation of cold, indicating the presence of cold spots. It will be noted that certain of the spots are much more reactive than others, and that those of cold are much the more numerous (see Fig. 205). Both heat and cold spots are most frequent at the nipples; then, in order, come the chest, the nose, the anterior surface of the arm, and the abdomen. They are least marked on the exposed surface of the skin, such as the face, and they are

also very infrequent in the scalp. They are almost absent from the mucous membranes, which explains why one is able to swallow a liquid that is too hot for the hand.

The *acuteness of the temperature sensation,* as with all the other cutaneous sensations, depends very much on the condition of the skin, being most sensitive when this is at the ordinary temperature, but very imperfect when it is either very hot or very cold. There is also very marked *adaptation of the sense.* This can be very well shown by the simple experiment of taking three vessels of water, one at a moderate temperature, one very hot and one very cold. If a finger of one hand is placed in the hot water and a finger of the other in the cold, and they are left there for a short time, until the skin has assumed the same temperature as the water, and then transferred to the lukewarm water, the finger



A. *B.*

Fig. 205.—Cold spots (*A*) and heat spots (*B*) of an area of skin of the right hand. In each case the most intense sensations were experienced in the black areas, less intense in the lined, and least in the dotted. The blank areas represent parts where no special sensation of either kind was experienced. (From Goldscheider.)

transferred from the cold water will feel hot, and that transferred from the hot water will feel cold. Temperature sensation also produces a marked *positive after-effect.* Thus, if a cold coin is placed on the forehead and then removed, the cold sensation will persist for some time in the area of skin on which the coin was laid.

That the receptors for heat and cold respond only to one kind of stimulus, or if to others, only when these are excessive, can be well illustrated by the experiment of touching a cold spot with a very hot object: the sensation will be that of cold. The hot object has so pronounced a power of stimulation that it has overstepped the threshold for heat of the cold-adapted receptors. The sensation of cold is elicited more promptly than that of warmth. The distinction between a warm and a

hot bath may really depend on the fact that in the latter the cold spots are stimulated as well as those of heat. It is at least interesting to note that the physiologic reflexes stimulated by either a cold or a very hot bath are the same; thus, a rise of blood pressure and a contraction of the muscles of the skin occur in both cases.

The Touch Sense.—In order to investigate the touch sense accurately, von Frey has devised a method of using hairs of different thickness each mounted on a different handle. The hair which produces a sensation of touch when pressed on the skin so that it just bends is then similarly pressed on one scale pan of a balance, and the weight required in the other scale pan to hold the beam horizontal when the hair just bends, is ascertained. From the diameter of the hair one can then calculate how many grams per square millimeter are necessary to elicit the sensation of touch. The following quantitative results have been obtained by applying von Frey's method to different parts of the body:

	Gm. per sq. mm.
Tongue and nose	2
Lip	2.5
Finger tip and forehead	3
Back of finger	5
Palm	7
Forearm	8
Back of hand	12
Calf, shoulder	16
Abdomen	26
Outside of thigh	26
Shin and sole	28
Back of forearm	33
Loin	48

That the sense of touch is located in spots—*touch spots*—can best be demonstrated on the calf of the leg. If this is shaved and then carefully explored with a fairly stiff hair, it will be found that there are only some twelve to fifteen spots in an area of a square centimeter at which the hair can be felt. Between these spots there is no sensation of touch. That these spots are composed of specialized receptors can be very clearly shown by pressing a fine needle into one of them, when no pain will be experienced but only a peculiar shotty sense of pressure.

Careful examination of the position of the touch spots will further show that they are grouped around hair follicles, particularly on the side from which the hair extends—the windward side, we may call it. This fact explains the well-known experience that an object may be felt more acutely on a hairy surface than after that surface has been shaved. The hairs bend slightly when the object comes in contact with them, thus

causing pressure to be exerted on the hair follicles, so that the touch corpuscles in the neighborhood of the follicles, or perhaps the fine nerve plexus which surrounds them, becomes excited. The influence of hairs in increasing the touch sensation can be demonstrated by the von Frey method; for example, in one experiment over an area of 9 square millimeters of skin with hairs present, 2 milligrams were found to produce the sensation, whereas after the hairs had been removed, it required 36 milligrams.

The frequency of touch corpuscles differs very much in different parts of the body. They are most plentiful on the fingers, relatively infrequent over the skin of the back, and very scarce in the skin directly over bony surfaces. They are entirely absent from the cornea, the conjunctiva of the upper lid, and the glans penis. The adequate stimulus for touch is evidently deformation of the surface. Pressure exerted over all the touch corpuscles of a portion of skin is not felt. This can be demonstrated by dipping the finger into mercury. The pressure of the mercury is felt on the surface but not in the submerged portion of the finger. Touch is the most responsive of all the sensations. Thus, as has already been noted, a tuning fork can be felt vibrating by the finger when to the ear its note is a continuous one, and the stimuli produced by a revolving serrated wheel can be felt by the fingers as separate even up to a rate of five or six hundred stimuli per second. Adaptation is also a marked feature of the touch sense, as is the experience of every one who has worn flannel underclothing or a plate of false teeth.

Closely related to the tactile sense is the power of *discrimination between two points*. This is tested by finding at what distance the two points of a pair of calipers stand in order to be distinguished as separate. The result in any given part of the body varies a little according to whether the points rest on touch corpuscles and according to the relationship of the calipers to the hair follicles. On an average, however, we may take the following distances in millimeters as being those at which the two points can be distinguished over different areas of the body:

	mm.
Tip of tongue	1.1
Volar surface of finger tip	2.3
Dorsal of first phalanx	6.8
Palm of hand	11.3
Back of hand	31.6
Back of neck	64.0
Middle of back, upper arm and side	67.1

It is clear from this list that the power of discrimination tends to diminish in proportion to the lessening mobility of the part. It is greatest

at the tip of the tongue and the tip of the fingers; it is least on the relatively immobile skin of the back. These distances are much less when the points rest on two touch corpuscles. Under these conditions, for instance, the distance for the volar side of the finger tip or even for the palm of the hand may be only one-tenth of a millimeter; and for the arm and back it may become reduced to half a millimeter.

Localization of touch is a very accurate process, at least in the most sensitive parts of the skin, but nevertheless it is very probably a matter of education. An evidence of this is the fact that in the much more highly specialized retina the power of localization of objects in the visual field is a process of education and experience. For this reason a person from whom a congenital cataract has been removed, can not locate the objects which he sees until after he has learned by his experience of touch, taste, etc., to associate the portion of the retina stimulated with a certain part of the visual field. If this is true for the retina, it is also probably true for touch. The famous experiment of Aristotle is explicable on the same basis. If the fingers are crossed and a marble placed between the crossed fingers, it will be felt as double, since now it touches two skin surfaces which have not been accustomed to touch the same object, but educated to feel different objects. Experience associates those two skin areas with different objects, not with the same object.

The Pain Sense.—It was at one time thought that the sensation of pain was due to overstimulation of any kind of receptor, but it is now known that for this, as for other skin sensations, there exist special receptors. Thus, it is found that in certain parts of the body, such as the cornea, and to a certain extent in the glans penis, pain receptors alone are present, and in disease the sense of pain may be entirely abolished, whereas that of touch remains, this condition being called analgesia. Overstimulation of a touch spot does not, as we have seen, cause pain but only a sense of pressure. Although pain is appreciated by special receptors, the character of the pain is dependent on the other sense receptors simultaneously excited; for example, a throbbing pain is due to the simultaneous pressure produced by dilated blood vessels, etc. A sensation of pain accompanies certain reflexes of a protective nature (nociceptive reflexes, page 825), and when the reflex is absent the part is likely to suffer damage. On this account the pain nerves may be regarded as trophic nerves. The sense of pain may also occur in structures which are devoid of ordinary sensibility, such as the intestine and the ureter.

CHAPTER LXXXVIII

THE PROPERTIES OF EACH PART OF THE REFLEX ARC (Cont'd)

THE NERVE NETWORK

In all animals above the Celenterates, no direct protoplasmic continuity exists between the various neurons, the transmission of the nerve impulse depending on contiguity rather than continuity of the elements that constitute the reflex arc. This transmission may be effected through a synapsis coming in contact either with dendrites or with nerve cells. It is extremely difficult to know whether there is really any anatomic continuity between the various fibers which form the network in the gray matter of the central nervous system. We shall not attempt to discuss this vexed question here, but in order that we may learn something of the possible functions of a nerve network, we may consider that present in the walls of the intestine (plexus of Auerbach and Meissner.) This plexus seems to have an important function to perform in connection with the myenteric reflex (see page 466). At least it has been shown by Meek[3] that after transsection of the intestine the muscular and epithelial structures become regenerated considerably earlier than the nervous plexus, but that the myenteric reflex, which, it will be remembered, is characterized by a wave of inhibition preceding one of contraction does not occur until after the plexus has been regenerated.

NETWORK ON SKIN NERVES

A very important type of nerve network, from the medical viewpoint, is that which is produced close to their receptor endings by the branching of the afferent fibers of the skin. Through these branches the vascular reactions following the application of an irritant to the sensory surface take place without the intervention of any nerve cells. It used to be thought that such reflex vasodilatation depended upon the transmission of an impulse along an afferent neuron to an efferent vasodilator neuron, a view strictly in consonance with the neuron hypothesis. That such is not the case, however, is shown by the fact observed by Ninian Bruce[4] that irritants such as mustard oil applied to the skin or cornea continue to produce their usual reaction for some time after

section of the posterior roots of the spinal cord, but fail to do so if the nerve fibers are cut and allowed to degenerate, or if the stimuli are blocked by applying cocaine to the skin. What actually happens is evidently that the impulse set up by the irritant as it travels up the afferent fiber passes on to one of the branches above referred to, along which it then proceeds to the blood vessels, which it causes to dilate. That such vasodilator impulses may be transmitted down the fibers of an *afferent* nerve has been confirmed by Bayliss, who found that vasodilatation occurred in the hind limb when the posterior spinal roots were stimulated (see page 234).

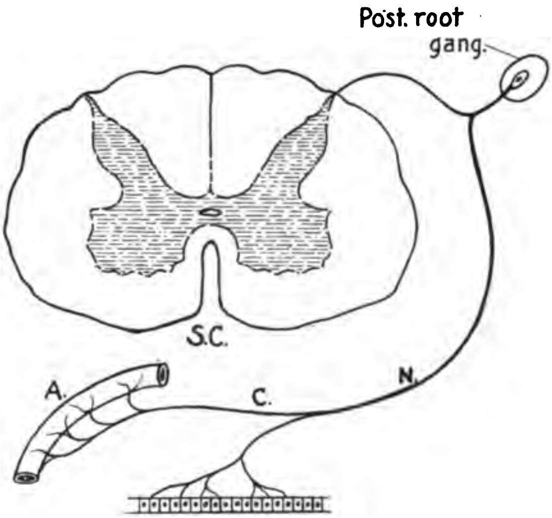


Fig. 206.—Diagram to show axon reflex of sensory nerve fiber of skin. A stimulus applied to the skin is transmitted by the sensory fiber (*N*), part of it going to the spinal cord (*SC*), and part of it passing by the collateral (*C*) to the arteriole (*A*), which it causes to dilate.

In this peripheral branching of the afferent fibers of the skin, we have therefore a sort of neuropile which, like that of certain forms of Celenterates (see page 782), is capable of serving as a pathway for the transmission of a sensory impulse to an effector organ without the intervention of nerve cells. Such a reflex is known as an axon reflex, and it is evident that it may occur through any fiber which gives off branches, one traveling to a sensory surface, the other to some effector organ, as occurs in the hypogastric nerves to the bladder (see page 883).

THE SYNAPSIS

At the point of contact between a branch of one neuron and a nerve cell of the next, we have seen that there exists a structure known as the synapsis. Although this is described by histologists as a tuft-like

branching of the end of the axon (Fig. 207), it may really consist of a sort of membrane—the synaptic membrane. It permits the nerve impulse to pass in one direction only, from synapsis to cell. Of what this membrane may be composed, we do not know, but there can be no doubt as to its great functional importance in connection with the integration of the central nervous system; for example, failure of an impulse to pass between two neurons may be due to retraction of the synaptic membrane from the cell, or to alteration in its permeability towards the nerve impulse, perhaps as a consequence of changes in surface

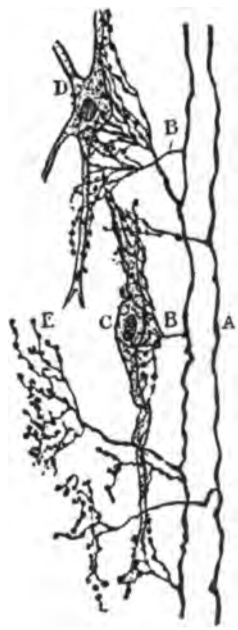


Fig. 207.—Arborization of collaterals from the posterior root fibers around the cells of the posterior horn. *A*, ascending fiber in posterior columns; *B*, collaterals; *C*, cells of posterior horn; *E*, synapsis. (From Ramon y Cajal.)

tension. Similar changes may also be brought about by the action of electrolytes or by chloroform, strychnine, and other drugs. As we shall see when we come to study the reflexes of the higher animals, there can be little doubt that it is in the synaptic membrane that many of the peculiarities reside which characterize conduction in a reflex arc as compared with that in a nerve trunk. The phenomena of summation, of reciprocal inhibition, of facilitation, etc., are undoubtedly dependent upon such alterations. The synapsis is also almost certainly the seat of fatigue in the central nervous system, and it is possibly the structure whose physiologic activity becomes upset in surgical shock.

THE NERVE CELL

Aside from being a meeting place of fibers coming from various sources, the nerve cell may have other functions, such as that of reinforcing impulses, just as a relay may reinforce an electric current. It is also responsible for maintaining the nutrition of the axon with which it is connected. In the case of the posterior root fibers of higher ani-

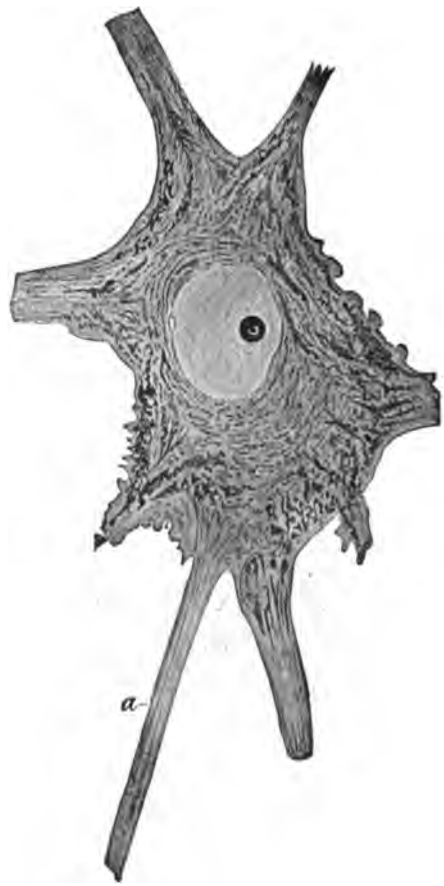


Fig. 208.—Normal cell from the anterior horn, stained to show Nissl's granules. *a*, the axon. (From Howell.)

mals, this function is probably the most important which the cell performs, for it has been found by separating the ganglia from their blood supply in the frog that, although the cells degenerate in about two weeks, sensory impulses continue to be transmitted through the ganglia. Similar observations have been made in the case of the crab, in which the cell bodies of the neurons lie on the surface of the ganglion

mass, from which they can be separated, leaving merely the neuropile, through which, however, the reflex continues to be conveyed. After a time, of course in this case also the reflex disappears, because an axon can not live indefinitely after it has been separated from its nerve cell.

These facts regarding the general function of the nerve cell arouse our curiosity as to its morphologic structure. When nerve cells are fixed and stained in various ways they show certain elements in the

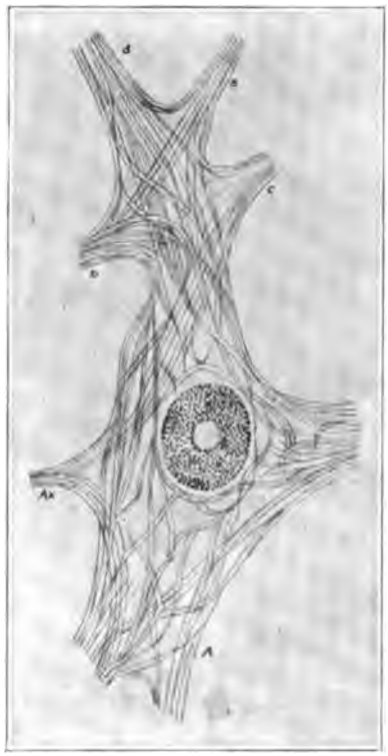


Fig. 209.—Part of an anterior cornual cell from the calf's spinal cord, stained to show neurofibrils. *ax*, axon; *a*, *b*, *c*, dendrites. (From Bethe.)

cytoplasm—namely, (1) large granules or masses, which stain deeply with basic dyes and are called Nissl bodies (Fig. 208), and (2) a fine network of fibrils passing through the cell substance from one process or dendrite to another—neurofibrils (Fig. 209). These appearances in fixed and stained preparations are possibly, however, entirely artificial; for when nerve cells are preserved in a living state—by being suspended in some of the animal's own lymph or blood serum—it is found, when they are examined by the aid of the ultramicroscope (see page 52), that the cytoplasm

is composed of a viscous fluid full of extremely minute granules, each of which apparently consists of a colloidal solution surrounded by a lipoid envelope (Fig. 210). When the temperature is raised, the granules disappear, and when the cells are deprived of oxygen, the cytoplasm and nucleus become swollen. A similar swelling of the cell and nucleus supervenes upon section of the axon; and in stained specimens the Nissl granules disappear and the protoplasm stains diffusely (chromatolysis).

In embryonic life the processes of the nerve cells appear to be capable of undergoing a certain amount of ameboid movement, and fibers grow out from them, indicating, therefore, that in the development of the nervous system the nerve cells appear first, and the nerves subsequently grow out from them to their proper destination. Proliferation of isolated tissue cells *in vitro* has been observed for many other

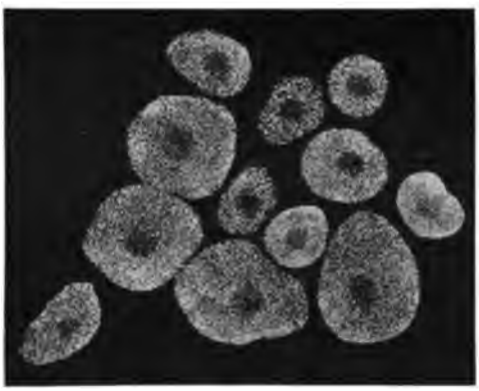


Fig. 210.—Living nerve cells (from the dorsal root ganglia of a dog three days old) examined by the ultramicroscope. There are no Nissl bodies or neurofibrils, only fine particles, present in the protoplasm. (From Marinesco.)

tissues, such as cardiac muscle, renal epithelium and connective tissue. Its occurrence indicates that the therapeutic principle that the aim of treatment should be to give the diseased organ a rest so that by cell regeneration it may recover its lost function, is one which may apply to the nerve tissues of young animals. Whether adult nerve cells may regenerate is as yet not certain.

This growing out of nerve fibers from their cells is the essential nature of the development of the nervous system in the developing animal. At birth, unlike the cells of other tissues, those of the central nervous system are already provided. No new ones are added during postnatal life. The axons gradually develop from this inherited stock of nerve cells, and by connecting with other neurons serve to bring about the integration which is the important characteristic of the adult nervous

system. The more complex the integration, the higher the intelligence of the animal.

Besides performing these functions the nerve cells serve as storehouses for memory impressions, certain types of them being especially adapted for this function. The differences observed in the relative thickness of the cell layers composing the cerebral cortex are more or less associated with the function which it can be shown the different areas of this possess. Nerve cells are extraordinarily sensitive to deficiency in oxygen supply, and yet little evidence of oxygen consumption by the brain can be revealed by the usual methods of investigation (page 396).

THE INTERMEDIATE OR INTERNUNCIAL NEURON

It would be profitless at this stage to consider the possible influences that the intermediate neuron may have on the impulses passing along the reflex arc. Before doing so we must see how the problem can be approached, for it is plain that the neuron in the case of the simpler reflexes is too short to make any investigation of its peculiar functions a possibility. We must study the characteristics of some type of reflex in which this neuron is drawn out, such as the scratch reflex, in which, as we shall see, it extends from the shoulder area of the cord to the lumbar region.

CHAPTER LXXXIX

THE REFLEXES OF THE SPINAL ANIMAL AND SPINAL SHOCK

Having become familiar with the peculiar properties of each of the structures which go to make up the reflex arc, we may now proceed to consider the function of the arc as a whole. It may be well first of all to consider briefly the experimental method by which such studies may be made. The object aimed at is to simplify the conditions as much as possible, for it will be evident that, in the intact nervous system, with the brain exercising a dominating influence over the great majority of all the reflexes, it would be impossible by applying a given stimulus, to predict exactly what kind of reflex response it might call forth. The reflex will be conditioned upon the accompanying influence which the brain exercises on the reflex involved.

In order to render the reflex unconditioned, we must remove the influence of higher centers. This can be done experimentally for the reflexes of a great part of the body by cutting the spinal cord above the level of the segment in which the reflex under investigation resides. Some of the reflexes elicitable from the cord isolated in this way involve only one or two neighboring segments, whereas others spread over several. The reflexes which have been most extensively employed are those which involve the musculature of the hind limbs. Since some of the receptors concerned come from the skin of the flank and shoulder areas, the section is usually made at the upper end of the thoracic region of the spinal cord.

Spinal Shock in Laboratory Animals

Immediately after the operation a profound condition of depression sets in, involving all the reflex arcs in the separated portion of cord. This condition is known as *spinal shock*. It supervenes in all classes of animals having a spinal cord, but is much more profound in the higher than in the lower animals. As a result of this depression, the part of the body below the section exists in a limp and flaccid condition, and the application of even very strong stimuli to the skin will evoke no form of reflex movement. In the case of the lower animals, such as the frog,

the condition begins to pass off in from twenty minutes to half an hour, after which a stimulus applied to the skin of the foot is followed by a typical flexion movement at knee and hip, the so-called flexion reflex. In the rabbit very little reflex response is elicitable for several hours after the operation, but in a few days the reflexes return completely below the level of the section. In the dog, on which a great deal of work has been done, the involved regions of the body are profoundly paralyzed. The skin is in a more or less unhealthy, unnatural condition, the surface cold, the hairs ruffled; and if care is not taken, the slightest abrasion of the surface may result in a nasty ulceration. On account of the paralysis of the centers of micturition and defecation, there is also incontinence of urine and of feces.

The Reflexes in the Spinal Animal

With reasonable attention, however, the dog makes a wonderful recovery. After an interval of two weeks the hind limbs, although completely paralyzed so far as voluntary movement is concerned, begin to show considerable signs of improvement. The first reflexes to return are those concerned with the deeper structures, such as the vascular reflexes, thus bringing the skin back to its normal temperature and condition. The reflexes of micturition and defecation also soon return, so that the animal no longer suffers from the continuous discharge of urine and feces. About the same time the *knee-jerk* becomes elicitable. This reflex is obtained by tapping the tendon which connects the patella with the tibia, the response being a smart contraction of the extensor muscles of the knee joint. The *flexion reflex* also begins to reappear. This is elicited by applying a pinprick or other hurtful stimulus to the skin of a lower extremity, and when fully developed consists in a flexion of the knee and hip joints. The evident object of this movement is that the stimulated parts may be removed from the source of stimulation, and it is plain that all stimuli that produce the flexion reflex are such as would cause in the intact animal a sensation of pain. Such stimuli are thus classified as nocuous, and the reflex is styled a *nociceptive reflex*. Accompanying flexion of the stimulated limb the opposite or contralateral limb usually undergoes a definite extension, called the *crossed extension reflex*. The occurrence of this together with the flexion of the stimulated limb is an important thing to remember in testing the reflexes in man. Malingerers who attempt to make it appear that they have some lesion of the spinal cord may know that if such lesion exists no movement of the leg occurs when the skin is stimulated, but they are unlikely to know that under these conditions the opposite leg also fails to show a simultaneous extension.

That the nociceptive reflexes should be among the first to return after spinal transection is of considerable interest as indicating their importance in the protection of the animal from injury. They are the essential reflexes of defense, and it is considerably later in the recovery of the animal before reflexes dependent upon stimulation of other tactile receptors begin to show themselves. The most important of this latter group of more special reflex movements include the so-called scratch reflex and the extensor thrust. The *scratch reflex,* as its name implies, is the scratching movement of flexion and extension of the hind limb at a rate of about four contractions per second that occurs when a mechanical stimulus is applied to the flank and shoulder area of the animal. For example, if we gently draw a pencil or the fingers backward and forward among the hairs on this region of the spinal animal, the corresponding hind limb will be brought up so that the claws are approximately at the place stimulated, and the limb thus directed will undergo a series of flexions and extensions, designed evidently for the purpose of scratching the area of skin that has been stimulated. If the stimulus is a weak one, only the initial stages of the movement may occur, such as the preliminary flexion of the leg. As we have already stated, the receptive stimulus calling forth this reflex is very specific in nature. A pinprick or rough friction of the reflex area will not produce it, nor will the application of heat or of a single electric shock. The most adequate stimulus is one simulating as nearly as possible the condition which would be produced by the movement on the flank of the animal of some insect. This more or less complicated scratch reflex can of course also be elicited in animals whose spinal cord has not been cut, but we can not predict in such cases whether the reflex will occur. The brain may inhibit the reflex arc and prevent the movement. In a spinal animal, however, the reflex always occurs provided an adequate stimulus is applied. The great importance of the scratch reflex in the study of the physiology of the spinal cord rests in the fact that a large stretch of cord is involved in the reflex path. The afferent impulses must enter at a much higher level than the efferent impulses leave, and between these two points there must exist a long intraspinal neuron (see page 813). This permits us to study many conditions influencing reflex action which otherwise in a reflex located in one segment only it would be impossible to investigate.[4a]

The extensor thrust is elicited by applying pressure to the pad of the paw or the sole of the foot. It consists of a quick extension movement of the corresponding limb usually with a flexion of the opposite limb.

After complete recovery from shock, the paralyzed parts of the body are capable of performing even more complex movements than those al-

ready mentioned. For example, if the animal is held up with the hind legs hanging down, these will often exhibit rhythmic flexion and extension movements, with the two limbs acting alternately, as they would in walking or running. This is sometimes called the ***mark-time reflex.*** Another complicated movement may be produced by placing the animal in water, when it may make the movements of swimming, but its swimming will not be sufficient to keep it on the surface. These swimming movements are more perfect in the spinal frog.

After complete recovery from spinal shock, the hind limbs are more or less in a condition of *extension contracture;* the vascular and other visceral reflexes are in perfect condition, and a marked rise in blood pressure occurs when one of the sensory nerves of the hind limb is stimulated—an experiment which can be performed in such animals without the administration of any anesthetic, since the animal feels no pain. In female spinal animals impregnation may occur and pregnancy proceed in normal fashion accompanied by the usual secretion of milk. The significance of this fact will be dwelt upon later.

SPINAL SHOCK IN MAN

As we ascend the animal scale we find that recovery from spinal shock takes longer and longer to occur and becomes less and less perfect. In the case of man, recovery is never complete, for a permanent condition, which has been called "isolation dystrophy," supervenes before the symptoms of shock have been recovered from. The tendon jerks are permanently abolished in complete lesions of the cord in man, and even when the lesion involves only one lateral half of the cord, this reflex is either entirely absent or very feeble on the corresponding side, though normal on the other (Holmes[5]). Severe lesions above the lower dorsal region practically always leave the legs in a permanently flaccid condition, with accompanying atrophy, but sometimes automatic movements of flexion and extension, like those of the mark-time reflex, may set in.

When the injury of the cord is less severe, the limb musculature is flaccid and toneless for some time, the tendon jerk and the abdominal and cremasteric skin reflexes being entirely absent. After some time, however—it may be as early as ten days—the muscles begin to reacquire some tone, and a little later the tendon jerk becomes elicitable. Regarding the behavior of the flexion reflex after spinal injuries in man, it has been found that the part of it known as the Bakinski reflex is not elicitable after severe lesions, but in those that are less severe a flexion of the great toe may occur on stimulation of the sole. Later this movement may be accompanied by contraction of the hamstrings, and later still, in favorable cases, by flexion at knee and hip. In these

cases also the Babinski reflex changes from a flexion to an extension of the great toe. It is important to note in connection with the above association of movements, that the sensory area of the sole is connected with the same segment of the spinal cord that furnishes the motor fibers to the flexors of the toes and the hamstrings (first sacral.) The recovery after shock therefore sets in earlier for unisegmental reflex areas than for those involving several segments.

The Cause of Spinal Shock

The relationship of the profundity of spinal shock to the phylogenetic position of the animal indicates that the shock must be due to the isolation of the lower spinal segments from the higher centers (Pike[6]). It has been suggested that the spinal section in the higher but not in the lower animals breaks a nervous pathway in which normally the reflex impulses travel. According to this view, the afferent impulse, when it enters the spinal cord in the lower animals, chooses the shortest possible route to the effector neuron of the same or closely adjacent segments by the collateral branches springing from the sensory neuron. In the higher animals, however, it would appear that, although this local spinal pathway is present and may be taken, yet it is usually passed by and the impulse travels up to the higher centers, from which it is then transmitted by the pyramidal tracts to the motor neurons concerned. This would appear to be the pathway for nervous reflex impulses in higher animals—the beaten track. When the spinal cord is severed, therefore, the condition of shock supervenes because impulses have not yet learned that they may find a shorter road to the motor neuron by the collateral than by the pathway which they usually travel. They learn this only after some time, which explains the slow recovery of the reflexes from shock.

It is obviously a difficult matter to supply direct proof in support of the above hypothesis of the cause of spinal shock, but besides the indirect evidence furnished by observations on the degree to which this condition supervenes in different groups of anmials, the hypothesis also conforms well with all the other facts which we know regarding the condition. For example, it is well known that the portion of the body above the transection of the spinal cord in no way suffers from the shock. Sherrington has described a monkey the cord of which was cut below the cervical region, and which immediately after the operation amused itself by catching flies with the anterior extremities, whereas the posterior extremities were in a condition of the profoundest shock. Such experiments further indicate that the shock can not be dependent

upon the lowering of arterial blood pressure which a section of the cord higher than the mid-dorsal region must entail. The poor nutritive condition of the skin which we have seen to exist in the hind limbs in shock, shows that the blood vessels in them are profoundly dilated, but evidently the fall in blood pressure has nothing to do with the faulty conduction through the spinal cord, for such a fall would affect the centers for the fore limbs as well as those for the hind, and yet the former show no symptoms of shock.

Exactly similar shock is obtained by any section of the spinal cord as high up as the medulla. Of course as the section is made higher and higher up, the resulting paralysis becomes more and more marked, and may reach such a degree of severity that recovery of the animal becomes an impossiblity.

When we come to consider the functions of the various parts of the brain, we shall have occasion to study the effects following section at higher levels of the cerebrospinal axis. Meanwhile, however, it is important to note that when a section is made across the crura cerebri, so that the cerebral hemispheres alone are isolated from the rest of the nervous system, a condition of contracture of all of the extensor muscles occurs. This condition is known as *decerebrate rigidity.*

CHAPTER XC

PHYSIOLOGICAL PROPERTIES OF THE SIMPLE REFLEX ARC

We may now proceed to study the properties of reflex action occurring through the isolated spinal centers of a spinal animal. There are two aspects of the question to be considered: (1) the properties of a single reflex arc, and (2) the action or influence of one reflex arc on another. The importance of the latter will be evident when it is remembered that complicated muscular movements depend for their proper coordination entirely on the interaction between the various reflex arcs which compose the nervous system. This interaction, as already explained, has been called by Sherrington the *integration of the nervous system.*

Probably the simplest way to study the physiologic properties of the simple reflex is to compare the mode of conduction of a nerve impulse through it with conduction along a simple nerve trunk. By comparing the two modes of conduction we shall be better able to appreciate the modifications to which the impulse is subjected by conduction through the reflex arc. The important points are these:

1. **The Latent Period.**—The latent period, or period which intervenes between the moment of application of the stimulus and the response, is very short in the case of a nerve trunk, and under normal conditions always the same, but is quite variable and sometimes very long in the case of a reflex arc. Thus, in the case of the conjunctival reflex, which is produced by applying a stimulus to the corneal conjunctiva (causing a closing of the eyelids), the reflex time is very short and invariable, whereas in the case of the scratch reflex it may vary from two and a half to three and a half seconds, according to the strength of the stimulus. The seat of delay in the reflex arc is probably in the synapse, but its cause is obscure.

2. **Grading of Intensity.**—In a nerve trunk the intensity of the impulse is more or less proportional to the strength of the stimulus. This can be judged by observing either the action current in the nerve by means of a galvanometer or the response of the end organ; e. g., muscle, attached to the nerve. In the case of a reflex arc, on the other hand, there is by no means so evident a parallelism between stimulus and response. Reflexes, however, vary considerably in this regard. The conjunctival reflex and the extensor thrust behave according to the so-called "all or nothing principle;" i. e., the intensity of the response is more or less independent of the strength of the stimulus. In other reflexes, such as the flexion reflex and the scratch reflex, the intensity of the response

is much more nearly proportional to the strength of the stimulus. Thus, a feeble stimulus applied to the flank calls forth only a slight flexion of the hind limb of the same side, whereas a stronger stimulus sets going a typical scratching movement.

3. **After-effect.**—When a stimulus is removed from a nerve, the effect which it produces, as judged, for example, by the action current, immediately disappears. There is no after-response. In reflex arcs, however, such a phenomenon is usually well marked. Particularly is this the case in the flexion and scratch reflexes of the spinal dog. A momentary stimulus of optimal strength applied to the scratch skin-area may produce no immediate response, but after its removal a violent scratching movement may set in. This after-discharge, in cases in which the stimulus is strong, may indeed, as in the flexion reflex, be more marked than the response during the time of application of the stimulus. In this particular reflex, the after-discharge often takes the form of a clonus, with a rate of contraction of from seven and a half to twelve per second. The crossed extension reflex also has a very pronounced after-discharge, which may outlast the stimulus for from ten to fifteen seconds. Regarding the phenomenon of after-discharge, Sherrington has stated that there is "no feature of the conduction of a reflex arc which distinguishes its mechanism more universally from that of a nerve fiber, tract or trunk than lengthy after-discharge."

4. **Summation.**—When a subliminal stimulus—that is, one that has in itself no visible effect—is frequently repeated in the case of a nerve, no response occurs. In the case of a reflex arc, however, such repetition of subliminal stimuli ultimately calls forth response. This summation is very evident in the case of the scratch reflex; e. g., one or two electrical stimuli applied to the scratch field-area call forth, as a rule, no movement of the corresponding hind leg, but if these same stimuli are frequently repeated, the typical reflex scratching movement will occur. Evidently, then, in a reflex arc there is a considerable amount of resistance towards a single stimulus, which resistance is overcome by a succession of stimuli. In other words, the threshold of the excitability of the reflex mechanism becomes lowered as a result of its previous stimulation. Each stimulus excites the sensory surface so that it responds more easily to the succeeding stimulus.

5. **Irreversibility of the Direction of Conduction.**—This is well illustrated in the so-called Bell-Magendie law of conduction in the spinal nerve roots. A motor impulse travels out of the cord by the anterior roots, while a sensory impulse travels in by the posterior. This directive influence can not depend on the nerve trunks or the nerve cells, for nerve trunks conduct equally in both directions, and so also must the nerve cell. The irreversibility must therefore depend on the synaptic

from the volitional center in the cerebrum. The same result may be obtained by electric stimulation of the center for eye movements on the cerebral cortex.

The most important details concerning the mechanism of reciprocal innervation have been obtained by studying the flexion reflex in a spinal animal which has completely recovered from shock. In such an animal the tonus of the extensor muscles of the knees is well marked. This tonus is maintained by afferent impulses transmitted to the spinal cord from receptors situated in the muscles, and its degree of intensity can be estimated by the briskness of the knee-jerk, which, it will be remem-

Fig. 212.—Record from myograph connected with the extensor muscle of the knee. During the time marked by the lower signal, the skin of the opposite foot was stimulated, thus causing the crossed extension reflex. While still maintaining this stimulation, faradic shocks were applied to the skin of the foot of the same side (as indicated by the upper signal), with the result that immediate inhibition of the contracted extensor occurred. (From Sherrington.)

bered, is elicited by tapping the patellar tendon, and consists of a sudden extension movement at the knee joint. By observing the briskness of the knee-jerk we are therefore enabled to form an estimate of the tonicity of the extensor muscles; and if after doing so we throw the flexors which are their antagonists into activity by eliciting the flexion reflex, the knee-jerk will be found much less active. If we prevent the flexors from acting on the knee joint and the leg is held in an extended position, irritation of the skin of the leg will cause the flexion of the

disconnected hamstring muscles simultaneously with a visible relaxation of the extensors (Fig. 212). If the leg is held properly, this relaxation may be marked enough to cause a slight flexion at the joint; and in any case, if the knee-jerk is regularly elicited by equal taps applied to the patellar tendon, it will be found that, while the flexion is being produced, the knee-jerk is very much less than normal, if not entirely absent, thus indicating that the tone of the extensor muscles is diminished. This experiment is very striking when performed on a decerebrate animal, in which, as we shall see, the extensor muscles of the limb are in a permanent state of hypertonicity (Fig. 213).

Before it is permissible to conclude that this reciprocal inhibition is a necessary event in the movement of a joint, we must however show that it occurs at exactly the same time as the flexion of the antagonist. Sherrington has succeeded in doing this in a considerable variety of experi-

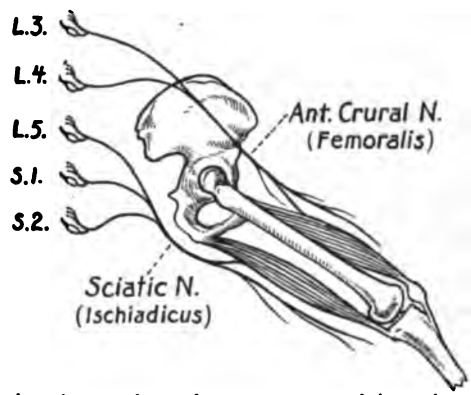


Fig. 213.—Diagram showing the muscles and nerves concerned in reciprocal innervation. (After Sherrington.)

ments, one of which we may cite here. If, in a spinal dog, the tendons of the flexor muscles of the knee joint of one hind limb and the extensor tendons of the opposite limb are cut, then the former limb will be unable to flex properly, but will nevertheless exhibit reciprocal inhibition of the intact extensor muscle, while the latter limb will flex, but require passive extension to bring it back to its old position. If suitable stimuli are simultaneously applied to the skin of both legs and the movements of the isolated muscles recorded, the onset of inhibition of the intact extensor of the one leg and the contraction of the flexors of the opposite leg will be found to agree with regard to latent periods, strength of required stimulus, summation and indeed all the other physiologic properties of reflex action.

Reciprocal innervation can also be demonstrated by stimulating the central end of suitable afferent nerves—that is, certain afferent nerves

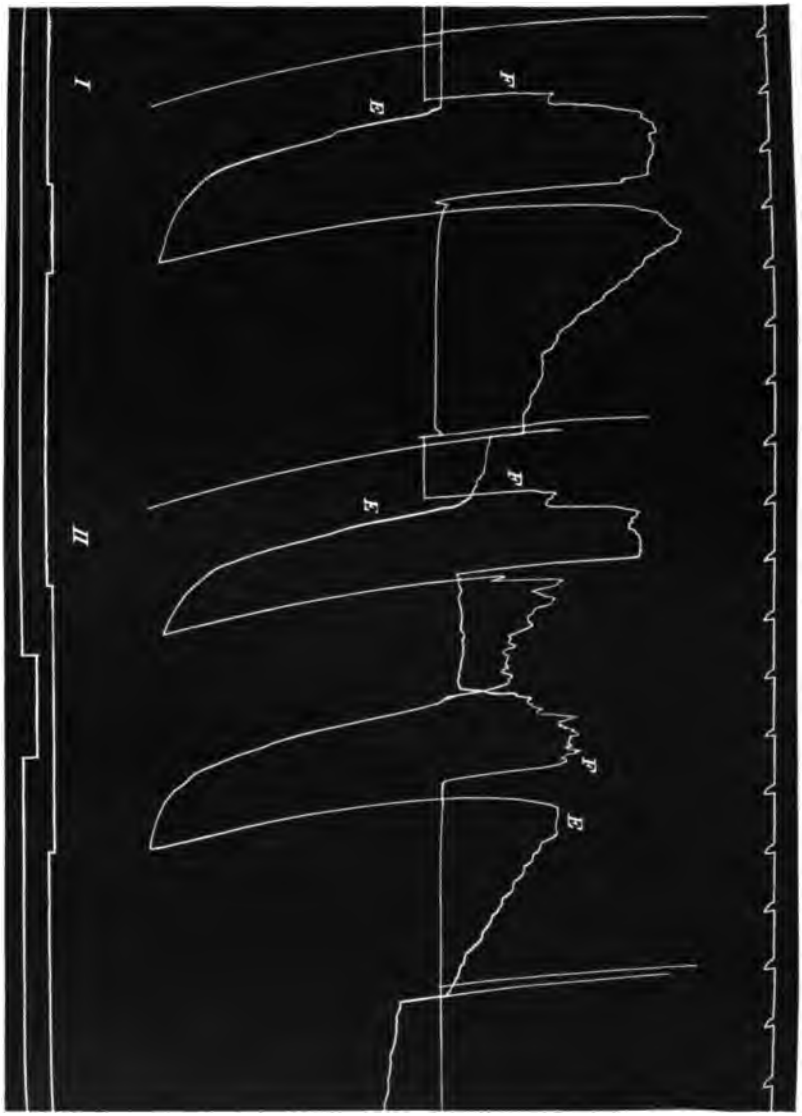


Fig. 214.—Reciprocal innervation. Tracings made by myographs connected with *E*, an extensor muscle (vastus crureus), and *F*, a flexor muscle (semitendinosus), of a decerebrate cat. At signal *I* the contralateral peroneal nerve was excited, causing contraction of the flexors and inhibition of the tone of the extensors. At signal *II* the flexors were again thrown into contraction by exciting the contralateral peroneal nerve, and (without removing this stimulus) the homolateral peroneal nerve was excited (as shown in the lower signal), with the result that the contraction of the flexors was inhibited at the same time that the extensors contracted. On removal of the latter stimulus, the former one reasserted its influence. This experiment demonstrates very clearly the accurate coincidence of the reciprocal action. (From Sherrington.)

acting on the same groups of neurons will produce a flexion reflex, others an extension reflex; thus, stimulation of the contralateral peroneal nerve produces a flexion reflex of the hind limb (excitatory for flexors, inhibitory for extensors), whereas stimulation of the homolateral peroneal nerve produces an extension (inhibitory for flexors, excitatory for extensors). By taking advantage of these facts further proof may be supplied that inhibition and contraction occur simultaneously, as shown in Fig. 214.

It is impossible to demonstrate any trace of inhibition of the skeletal

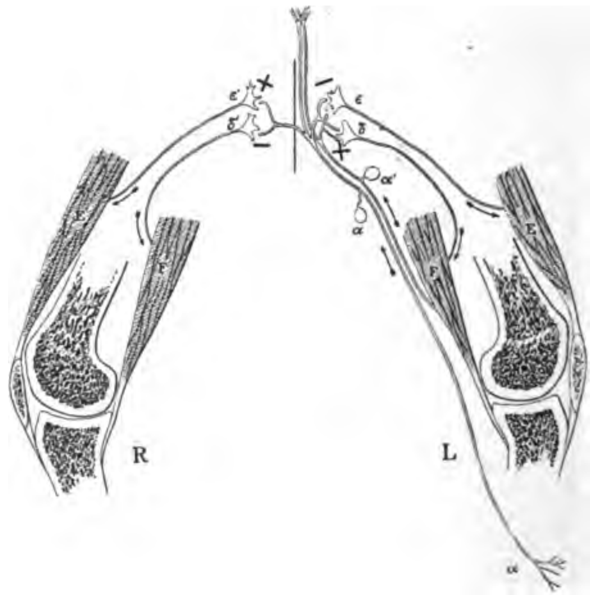


Fig. 215.—Sherrington's diagram illustrating the mechanism of reciprocal innervation. The afferent fibers (δ) from the skin of the leg and (δ') from the flexor muscles of the knee (in hamstring nerve) pass to the spinal cord, where each gives off a branch which divides into two others, of which one in each case goes to a motor neuron of the extensor muscles (E) and the other to a motor neuron (δ) of the flexor muscles (F). Branches also pass across the median line to similar motor neurons on the opposite side of the cord. As indicated by the plus and minus signs, the afferent stimuli either stimulate or inhibit the activities of the motor neurons, the determination of the exact effect being a function of the synapsis. (From Sherrington.)

muscles by stimulation of their motor nerves, thus indicating that inhibition is dependent upon the nerve center. Furthermore, since inhibition occurs along with flexion of the antagonistic muscle, we must assume that the afferent impulse on entering the spinal cord divides into two branches, one going to one motor neuron so as to excite it, the other to another neuron so as to inhibit the tonic stimuli which it is constantly sending to the muscles (Fig. 215).

Since the seat of the inhibition is in the nerve center, it is to be expected that impulses transmitted from other parts of the nervous system

than the particular level of that reflex, will also be able to induce the inhibition. In the case of the decerebrate cat this can be demonstrated by stimulation of the lateral columns of the spinal cord; inhibition of the extensor muscles of the elbow joint occurs, which is all the more marked because in such a preparation these muscles are in a state of hypertonicity. We shall see later also that through the pyramidal tract impulses may descend from the cerebrum which exercise a marked inhibitory influence over the reflex activities of the cord. Similarly the inhibition itself may be terminated by impulses from other sources, and the motor neuron thus thrown from a state of inhibition into one of excitation. This fact can perhaps best be demonstrated by exciting the central end of the homolateral peroneal nerve (which produces a reflex extension of the leg) while the leg is being held in a flexed position by stimulation of the contralateral peroneal nerve. This will be clear from a study of Fig. 214.

Such alternating excitation and inhibition of an active motor neuron serve to make it possible for rhythmic discharges to occur through the neuron, as in the action of the muscles of the leg in walking or during the scratching movement. In order to insure that the same final common path may be occupied at one time by but one kind of stimulus, either inhibitory or excitatory, it is further of importance that the after-discharge (see page 810) of the first stimulus should be capable of immediate inhibition; otherwise, while one reflex was in progress, it would be impossible to start another of a different type employing the same motor neuron without confusion of movement. That this occurs can be demonstrated in the case of the after-discharge of the flexion reflex by stimulation of the proper afferent nerve.

In view of all these facts it is probable that the seat of the reciprocal innervation is at or about the synapsis. In other words, the synapsis at the termination of one collateral will allow a stimulating impulse to pass to the cells of one motor neuron, whereas that at the end of another collateral of the same afferent fiber will allow an inhibiting impulse to pass to an antagonistic motor neuron, these conditions being, however, readily interchangeable and thus making even rapid rhythmic contraction and relaxation a possibility.

The Action of Strychnine and Tetanus Toxin on Reciprocal Inhibition

Under certain conditions reciprocal action may fail to occur, as, for example, at certain stages of *strychnine poisoning* and during the action of *tetanus toxin*. In order to demonstrate this failure of reciprocal action, it is necessary to examine muscles which act on one joint only, and

to observe their behavior when an afferent nerve is stimulated which under ordinary conditions would throw them into inhibition. Such a preparation can be obtained in the hind limb of a dog by cutting all the muscles that act on the knee joint except the vastus crureus, which in a normal animal invariably undergoes inhibition when the central end of the internal saphenous nerve is stimulated. If a suitable dose of strychnine is injected, it will be found that stimulation of the internal saphenous nerve, in place of inhibition, causes contraction of the vastus crureus muscle. The same result is obtained by injection of tetanus toxin.

The failure of the reflex inhibition explains the symptoms produced by these substances. It explains, for example, the well-known rigidly extended condition of the limbs in strychnine poisoning, and the distressing symptom of lockjaw in tetanus infection. In this latter condition the sufferer is subjected to extreme torture with every endeavor that he makes to open the jaw for the purpose of taking food or drink. Firmer closure is the result because the normal inhibition of the temporal and masseter muscles does not occur, but instead they become excited and the jaw all the more firmly closed. Not only does the inhibition fail to occur, but the above muscles are usually in a state of constant hyperexcitability, which it is impossible for the patient to restrain; indeed, whenever he attempts to do so the opposite occurs and the excitation becomes heightened. Chloroform acts on reciprocal innervation in an opposite way from strychnine and tetanus; namely, it paralyzes the excitation of the contracting muscles.

Finally, it must be pointed out that this mechanism of reciprocal innervation is by no means confined to the voluntary muscles. We have already seen that it occurs in the case of the myenteric reflex. It is also a most important function in the innervation of the blood vessels, dilatation in one vascular area being accompanied by constriction in another. These facts have been already sufficiently dwelt upon elsewhere (page 243). Sometimes also we may have reciprocal action between differently acting nervous mechanisms, as for example in the case of the submaxillary glands, which respond to stimulation of the chorda tympani nerve by dilatation of the blood vessels, an inhibition of their tone occurring along with stimulation of the activity of the gland cells.

CHAPTER XCII

INTERACTION AMONG REFLEXES

A single reflex acting independently of the rest of the central nervous system does not really occur. An afferent impulse on entering the cord spreads so as to involve a large variety of motor neurons, each of which may, however, be excited through other afferent fibers arriving either from other receptors or from higher nerve centers. The motor neuron itself may therefore be a pathway occupied at different times by very different types of nerve impulse. Hence it is appropriately called *the final common path*, and its activity at any moment must depend on the nature of the various afferent impulses that are transmitted to it through the synapses. In other words, an entering afferent fiber must communicate in the cord with internuncial paths which are available in various degrees to other afferent fibers. Since it is through internuncial paths that the impulse is transmitted to the final common path, it is obvious that, if afferent impulses in several of these paths were competing at the same time for the possession of the final common path, confusion of movement would result unless some provision were made whereby only one kind of stimulus could be transmitted at one time. "One kind of stimulus must be inhibited and the other facilitated in its occupancy of the final common path."

To understand the nature of this integration of the central nervous system, it is therefore necessary for us to consider the factors which determine *which of two competing afferent impulses shall obtain possession of the final common path.* Let us take the competition between the flexion reflex and the scratch reflex of the spinal dog. If we elicit the scratch reflex and, while it is in progress, apply some nocuous stimulus to the skin of the hind leg and thus induce the flexion reflex, it will be found that the scratching movement subsides and the flexion movement comes on without any overlapping or confusion. If, however, the stimulus responsible for the scratching movement is a strong one, and that applied to the skin of the hind leg a feeble one, then the displacement may not occur (see Fig. 216).

In considering this integration of reflexes, as it is called, we must distinguish between those that are allied and those that are antagonistic, and we must further distinguish between reflexes that are simultane-

ously competing for the same final common path and those which occupy it successively.

INTEGRATION OF ALLIED REFLEXES

Perhaps the simplest experiment to show this is performed by using the scratch reflex. The skin area from which this reflex can be elicited is very widespread (see Fig. 217), the type of reflex produced from any given area being in general the same, although "the local sign"—that is, the point at which the animal scratches—will vary according to the point stimulated. If then we take point *A* in the reflex scratch area and apply to it a stimulus which is just inadequate to produce any reflex at all, and then, while this stimulus is still in progress, apply a similar subliminal stimulus to point *B* a little removed from it, the two sub-

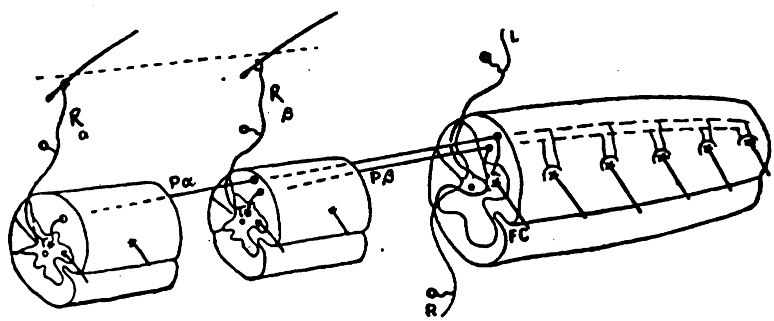


Fig. 216.—Diagram showing the reflex arcs involved in the scratch reflex. $R\alpha$ and $R\beta$ represent the afferent neurons connected with hairs on the skin of the back and flank. The afferent impulses are transmitted by these fibers, and on entering the corresponding segments of the spinal cord terminate by synapses on cells of the internuncial neurons, whose arrows $P\alpha$ and $P\beta$ travel down in the lateral columns to terminate similarly around the cells of the motor neurons that innervate the muscles of the hind limb. Since afferent impulses coming from elsewhere, particularly from the skin of the leg (*R* and *L*), also terminate on these neurons and may excite them to a different type of action, the motor neuron is called the final common path (*F.C.*). (From Sherrington.)

liminal stimuli will become effective and produce a typical scratching movement. In other words, the subliminal stimulus of point *A* becomes added on the final common path with the subliminal stimulus of point *B;* the one has reinforced the other and produced, therefore, a *simultaneous integration of allied reflexes.*

The receptors from which these mutually reinforcing impulses are received need not, as in the above example, be of the same kind, similar results being obtained by stimulation of receptors of widely different kinds, such as exteroceptors and proprioceptors (see page 788). For example, if a stimulus inadequate to elicit a flexion reflex is applied to the skin of the leg, and another stimulus, itself also inadequate, is applied to the central end of some deep afferent nerve in the same leg, then the two subliminal stimuli will become effective in producing a

flexion movement. Nevertheless, the more closely allied the receptors are to one another, the more easily does summation occur.

The mutual reinforcement of allied reflexes lasts for a short time after the stimulation has been removed, the phenomenon being now known as *successive integration of allied reflexes.* It can be illustrated also in the case of the scratch reflex. If point *A* on the skin area is excited with a stimulus that in itself would be inadequate, immediately after an effective stimulus has been discontinued at point *B*, then the scratch movement will be kept up smoothly although it will of course become modified in local sign. For the same reason, a moving stimulus applied to the scratch area is far more effective than a stationary stimulus applied over the same extent of area. In such a case the stimulus that excites a reflex tends by its occupancy of the nervous pathway to facilitate the spread along the same pathway of succeeding allied stimuli; towards such it lowers the threshold of excitability of the reflex arc.

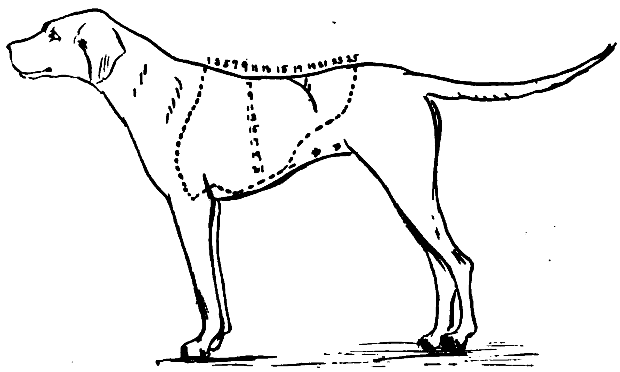


Fig. 217.—Showing region of body of dog from which the scratch reflex can be elicited. (From Sherrington.)

This phenomenon is also often called *immediate induction,* and it is by no means confined to the spinal cord. It is well illustrated, for example, in the case of vision. If a thin line drawn on a white card be looked at so that it falls on the edge of the receptive field of the retina, it will not be seen so well as a dot of similar width which is moved through the same distance as the line.

From these facts we see, therefore, that, when two allied impulses are being transmitted to the final common path, the one is likely to reinforce the other, and that this tendency to reinforce the allied impulse is maintained for a brief period of time after the impulse has been removed. We may now proceed to consider the factors which will become operative in determining to which of two competing or antagonistic reflexes the final common path will become available.

Integration of Antagonistic Reflexes.—Although the phenomenon of immediate induction encourages integration of allied reflexes, yet it is frequently succeeded by one of *successive induction*, in which just the opposite conditions occur; the resistance in the reflex pathway becomes lowered for a type of movement antagonistic to that which first occupied the reflex. To understand clearly what relationship this bears to immediate induction, it may be well to take the instances in which these phenomena apply in the case of vision. When the eye, after darkness, is suddenly directed to a light and then closed, there remains a bright image (positive after-effect) of the light; but if the light is looked at for some time, then on closing the eyes it will be seen as a dark pattern (negative after-effect). In the former instance we have an example of immediate induction, in the latter, one of successive induction.

In the spinal animal successive induction is demonstrated with equal ease by using two reflexes that are of a more or less antagonistic character—for example, the flexion reflex and the knee-jerk, or better still the crossed extension reflex and the flexion reflex. If we elicit the knee-jerk in a spinal dog at regular intervals, with stimuli of equal intensity, the extension movements (the kicks) will be approximately equal. If now we apply a nocuous stimulus to the skin of the foot and so throw the leg into flexion, it will be found, after the flexion movement has disappeared, that the knee-jerk is much more pronounced than previously. Similarly, if we elicit the crossed extension reflex by nocuous stimuli of equal intensity applied to the opposite limb, the extension movements will be approximately equal. By now throwing the limb exhibiting them into the flexion reflex, the extensor movements will of course disappear, but after the flexion has been discontinued, they will reappear with marked intensity.

These facts show us, then, that after the final common path has been occupied by a reflex of one type, it becomes more available to a reflex of an opposite type. In other words, it is evident that if the two opposite reflexes are constantly competing with each other for possession of the final common path, they will tend alternately to occupy it, thus bringing about a rhythmic movement. Such is the mechanism involved in walking: the leg is lifted from the ground (flexion reflex); it is then brought on the ground, and the mechanical push given to the plantar surface of the foot brings out the extensor thrust, the appearance of which is greatly facilitated by the fact that immediately before the flexion reflex occupied the final common path.

Other Factors Which Determine the Occupancy of the Final Common Path.—Besides immediate and successive induction, several other factors affect the relative availability of the reflexes to afferent stimulation.

to the tendon and the electrical response observed in the vastus muscle by the string galvanometer, was found by Jolly in the s to be 0.0055 of a second, whereas measured in the same way t period of the flexion reflex was found to be just twice as long; i. e., a second. These differences were explained by Jolly as indica the knee-jerk is a simple reflex, involving but two neurons, wh flexion reflex involves three and therefore has twice as long period. By subtracting from the total latent period the time in the transmission of the impulse along the nerves and the tin the afferent and efferent nerve endings, we secure a figure giving lost in the synapses between the neurons. This synapse time called, was found by Jolly to be 0.0021 of a second for the l and 0.0043 of a second for the flexion reflex.[7] Snyder obtained similar results in man by the same method.

Some authors, particularly Gowers, do not, however, believe knee-jerk is of the nature of a simple reflex, but explain it as b to a contraction of the extensor muscles brought about b mechanical stimulation while the muscle is in a hyperexcitable as a result of a reflex increase in its tonicity. Gowers believes putting the extensor muscles on the stretch and the hamstring in the relaxed condition, afferent impulses are transmitted to which excite the efferent neurons of the extensor muscles, so as them into a hypertonic condition, during which the tapping of don directly excites a contraction. Of course this hypothesis w count once and for all for the remarkably short latency of t jerk, but on the other hand it leaves us many difficulties to such, for example, as the fact that, although tapping the tendon the jerk, similar tapping of the muscle itself has no effect.

The effective stimulus of the jerk is a slight passive increa tension to which the extensor muscle itself is subjected, and not lation of receptors in the tendon, for it still occurs after the te been denervated. The importance of the relationship of the h nerve to the knee-jerk becomes evident in connection with r action; thus, when the flexor is contracted, as in the flexion r knee-jerk disappears (page 814), whereas when the hamstring n cut, it is augmented.

Whatever its nature may be, the knee-jerk is of value becau ease with which it can be altered not only by conditions affe reflex arc concerned, but also by changes occurring elsewhe central nervous system. The best known of these condition known as *reinforcement*. This is brought about by having th make some voluntary muscular effort at the moment that the

[redacted]

plied to the tendon. If this voluntary effort coincides in time with the tapping of the tendon, the knee-jerk will be found much augmented; but if the two events do not accurately coincide, we may find instead that the knee-jerk is diminished; that is to say, we may have positive followed by negative reinforcement. The most usual way of having the patient make this voluntary effort is to ask him to lock the fingers of his two hands together and then at a given signal try to pull the locked arms apart.

Similar reinforcement may also be produced by the application of a strong sensory stimulus in some distant part of the nervous system, as, for example, by pulling the hair or pinching the ear. Accurate work on the time relationship between the reinforcing act and the tap on the tendon has shown that the knee-jerk is most marked when the tap accurately corresponds with the voluntary effort or sensory stimulation. It then quickly declines and an inhibitory influence appears in about 0.3 to 0.6 of a second, immediately after which it becomes pronounced again, gradually fading off to be no longer evident in about 1.7 of a second; that is, no change from the normal will be found in the knee-jerk in about 1.5 of a second after the reinforcing act (Lombard[8]).

Many explanations have been offered of the mechanism involved in this reinforcement. The most commonly accepted is that it is due to the overflow of impulses from other parts of the nervous system, particularly the cerebrum, upon the reflex arc concerned in the knee-jerk. During voluntary effort the cerebral impulses discharged down the spinal cord pass not only to the neuron for which they are intended, but irradiate or spread to other, even far distant, neurons, thus adding their effect to that of the afferent impulse entering the cord locally. The succeeding inhibition may be assumed to be due to successive induction (see page 824). It is difficult to offer direct experimental proof in support of the explanation, but indirect evidence is furnished, in so far at least as the augmentation is concerned, by the results of the experiments which we have already described concerning the integration of allied reflexes (page 822). To these might be added the well-known fact that the simultaneous application of two subliminal stimuli, one to the cerebral cortex and the other to the skin of the corresponding body area, may call forth a contraction of certain groups of muscles.

AFFERENT SPINAL PATHWAYS

The nature of the impulses transmitted by the various afferent pathways in the spinal cord. We have seen that the sensory impulses traveling from the periphery to the spinal cord group themselves into three

CHAPTER XCIV

EFFECTS OF EXPERIMENTAL LESIONS OF VARIOUS PARTS OF THE NERVOUS SYSTEM

Having learned the main characteristics of reflex action, we shall now proceed to study the peculiar function of each part of the cerebrospinal system by noting the effects which follow destruction or stimulation of its different parts.

THE ANTERIOR ROOT

Section of an anterior root produces a limited degree of paralysis involving several muscles having no functional relationships to one another. If several anterior roots are cut, the paralysis becomes much more extended, and is followed very soon by an evident atrophy of the muscles concerned. Reflex actions from these muscles are of course impossible. Stimulation of the peripheral end of a cut motor root causes partial contraction of several muscles, no definite joint movement, however, being the result, because the affected muscles are not functionally related and there is no reciprocal inhibition. Flexor and extensor, adductor and abductor may contract at the same time, thus causing the joint on which they act to become muscle-bound. It is in the plexus that the nerve fibers of the roots become sorted out, according to function, into motor and sensory nerve trunks. The distribution of the anterior root fibers according to segments in man for the cervical and lumbosacral regions is as follows:

C5	Deltoid, biceps, brachialis, supinators, rhomboids. Occasionally radial extensors. Rarely pronator radii teres.
C6	Pronators, radial extensors, pectoralis major (clavicular fibers), serratus anticus.
C7	Triceps, extensor carpi ulnaris, extensors of fingers, pectoralis major.
C8	Flexors of wrist and fingers.
T1	Intrinsic muscles of hand.
S3, 4	Levator ani, sphincter ani, perineal muscles.
S2	Glutei, biceps, semitendinosus and semimembranosus.
S1	Intrinsic muscles of foot, tibialis posticus, and muscles of calf.
L5	Muscles of ventrolateral leg (except tibialis anticus).
L4	Extensors of leg and tibialis anticus.

The knowledge of the segmental innervation of the limb muscles, as furnished in the above table, is of value in the localization of spinal lesions. Paralysis of the extension movements of the wrist and fingers, along with the triceps, for example, usually indicates a lesion of the seventh cervical. It is more particularly in the trunk, however, that the segmental innervation of the muscles is evident. The innervation of the intercostal muscles being unsegmental, one may diagnose the level of a lesion of the upper thoracic region of the cord by observing their behavior during deep inspiration. If the fingers are placed in the intercostal spaces, the paralyzed muscles will feel limp and the fingers sink into the space during the act.

Localization may also be shown by studying the paralyses of the abdominal muscles when the lesion involves one of the lower six thoracic segments. When the patient with a lesion of the eleventh thoracic raises his head from the bed or coughs, the rectus contracts, but the iliac regions bulge owing to paralysis of the lower portions of the obliques. Under the same conditions, when the ninth segment is involved the rectus contracts from about one inch above the umbilicus, whereas below this level it remains uncontracted, so that the umbilicus is pulled up.

Besides muscular movement, stimulation of the anterior roots in lightly anesthetized animals sometimes causes evidence of general reflex response and of pain. The explanation is that there are present in the anterior root certain sensory fibers which are derived from the posterior root but recur in the anterior, so as to reach the membranes of the spinal cord where they terminate. The stimulation of the peripheral end of the motor root must produce, therefore, the same reflex responses as stimulation of the central end of the sensory root. Stimulation of the central end of a motor root has of course no effect.

THE POSTERIOR ROOT

The *posterior root* is the pathway by which impulses of the various receptors enter the spinal cord. Section of any considerable number of posterior roots causes therefore, anesthesia of the corresponding skin and muscle areas, but such a result does not become evident when one root alone is cut, because the sensory area supplied by each root overlaps at least half of that supplied by the neighboring roots. Although it is often difficult to distinguish the segmental distribution in the ramification of the fibers of the motor roots by finding what muscles they influence, this is more evident in the case of the sensory roots. On the trunk itself this segmental arrangement is very plain, but in the

extremities it is not at first sight so clear, although it can be accurately worked out, as is indicated in the accompanying diagram (Fig. 218).

In attempting to determine the level of a lesion from the sensory paralysis, some confusion often arises on account of the oblique course of the decussation of the sensory fibers in the spinal cord, fibers for the different sensations not crossing at the same levels. For example, the appreciation of moderate temperature is often lost slightly higher than that of pain. The appreciation of the vibrations caused by drawing the base of a

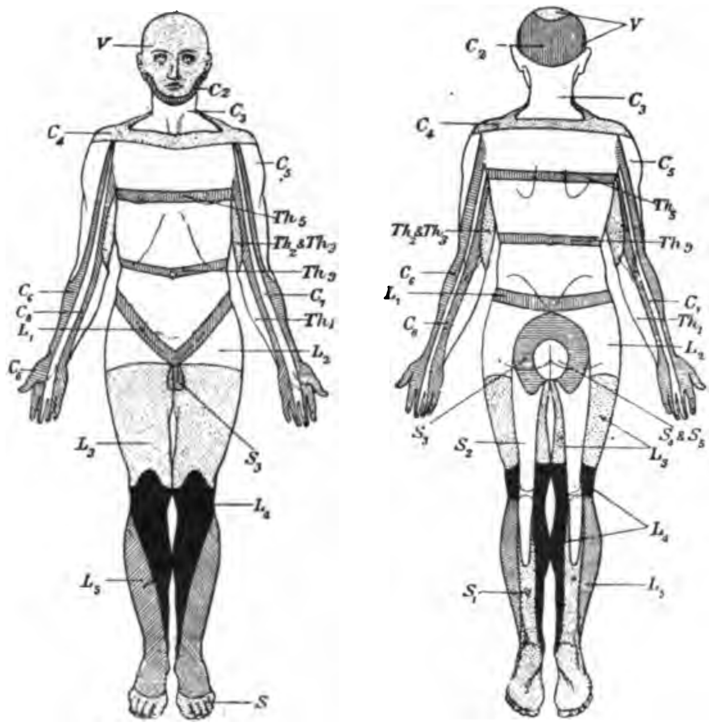


Fig. 218.—Diagram showing the segmental arrangement of the sensory nerves. (From Purves Stewart.)

heavy tuning fork over the skin is often very useful in locating the lesion, particularly in the abdomen. When this method is used on the thorax, however, the skin should be pulled up in folds before the fork is applied, since otherwise the thorax will act as a resonator and spread the sensation. Section of two or more sensory roots produces a very definite area of anesthesia, involving all the skin sensations as well as several of those of deep sensation.

If the severed roots include all of those going to one of the extremities, there is not only an entire absence of sensation, but a marked interference

with the usefulness of the limb, the condition being called ***apesthesia.*** The exact results depend somewhat on the type of animal. If all the posterior roots of the anterior extremity are cut in a ***monkey,*** the corresponding limb will not be used in climbing or for other purposes. It will appear to be completely paralyzed, unless when the opposite normal limb is in vigorous activity, when the apesthetic arm may be moved in association.

On careful examination, however, it will be found that marked differences exist in the types of paralysis produced by the section of the anterior and the posterior roots. When a motor root is cut no reflexes are possible either from the skin or from the cerebral cortex, and the muscles undergo atrophy. After section of the posterior root, on the other hand, although reflexes from the skin area affected are impossible, yet movements may be elicited by artificial stimulation of the cerebral cortex, and the muscles do not atrophy to the same extent.

If only one sensory root of an extremity is left uncut—for example, the last cervical—so that the skin of the hand is still supplied with sensation but all the deep receptors are severed, then the limb may be used to a modified degree. It may be used by the monkey to pick up nuts, but the movement will be distinctly clumsy and ataxic in nature. Instead of neatly picking up the nuts, he will make wild movements and often miss them.

The apesthesia is not so profound in lower animals. After section of all the sensory roots to both hind limbs in the *dog,* there may be a certain attempt at walking on the part of the affected limb; that is to say, when the animal tries to progress, the hind limbs, although at first merely dragged along the ground, afterwards begin to execute walking movements, which however are very jerky or ataxic in nature and contribute little to the forward progression of the animal, although he may succeed to a certain extent in supporting the body by the hind limbs.

The importance of the sensory root in controlling the contraction of the muscles is further illustrated by comparing the contraction curve of a muscle produced by stimulating its uncut motor nerve with that produced by stimulating the peripheral end of the cut nerve. In the former case, the curve is more prolonged and shows a gradual relaxation, whereas when the peripheral end of the cut nerve is stimulated, the contraction is brief and the relaxation is followed by a distinct rebound or "inertia swing," as it is called. That this difference depends on afferent impulses is indicated by the fact that, after section of the posterior roots, stimulation of the uncut nerve in the limb will produce the same effect as occurs when the cut nerve is stimulated. These results can be very clearly obtained in the case of the frog, in which

also it will be noted that after section of the posterior roots of one side, the corresponding limb hangs lower than its fellow because its muscles are toneless.

Stimulation of the central end of a cut afferent root produces, as has already been indicated, a contraction of the muscles accompanied by a reciprocal inhibition of their antagonists, so that some definite movement of the joint takes place. This movement is, however, merely a flexion or extension or rotation, but with no very evident object in view. In this regard it is quite different from the purposeful movement which results from stimulation of a skin area, indicating, therefore, that the receptor apparatus itself must contribute to the nerve impulse something which causes it to bring about a more perfectly integrated movement of the musculature than is the case when the nerve trunk is directly stimulated.

Besides the movements of the musculature innervated from segments which are beside those of the stimulated afferent root, there is a general reflex response through other centers, for example, the respiratory and the vasomotor; and, in animals which are not deeply anesthetized, there is also evidence of pain. Stimulation of the peripheral end of the sensory root has of course no effect.

THE SPINAL CORD AND BRAIN STEM

The results of transsection of the cord have been already sufficiently described. It remains to discuss the effect of **total ablation or removal of portions of the cord.** As would be expected, there is a marked degree of shock for some weeks after ablation. During this shock the tone of the sphincters and vessels is greatly depressed, so that congestion and edema of the feet, diarrhea and retentio urinæ are marked, and ulceration of the skin is practically unavoidable. After a few weeks, however, recovery becomes evident in so far as the blood vessels and sphincters are concerned, but the skeletal musculature atrophies very extensively and comes to resemble connective tissue. If the spinal ablation involves the thoracic region, for example, the affected intercostal muscles become stiff and parchment-like; the bones also get brittle, and visible perspiration can not be produced. On the other hand, after some time the sphincters functionate more or less normally, the hair is shed and renewed in normal fashion, and the application of cold to the skin causes the usual vascular reaction. It is of interest that in female animals whose lumbar spinal cord has been removed, pregnancy may take place normally, followed by lactation.

Section Just Above the Medulla.—After such an operation, the ani-

mal—bulbospinal, as it is called—shows a greater integration of reflexes than is possible when the section is between the medulla and the spinal cord. Its reflex responses are more broadly integrated, but the extremities are incapable of executing movements that are of any value in progression. Movements like those of progression may occur, but they are quite ineffective. Such animals show marked superiority over strictly spinal ones on account of the fact that in the medulla are located so many of the important centers which control circulation, respiration and the anterior openings of the body; that is, the mechanisms which accompany the first stages in the digestion of food.

Section Just Behind the Posterior Corpora Quadrigemina.—A very distinct improvement becomes noticeable in the responses of the animal. This condition has been studied most carefully in the case of the frog, which after such a section can walk, spring and swim apparently like a normal animal, and croaks when the side of the body is stroked. In the mammal a similar increase in the complexity of movement is evident, but there is not yet any power of progression.

Section in Front of the Anterior Corpora Quadrigemina.—When the medulla, pons and mesencephalon are present, as well as the spinal cord, the condition known as *decerebrate rigidity* supervenes. This is most marked in mammals, but is also present to a certain extent in much lower animals, as, for example, in frogs. It consists of a tonic condition of the postural musculature of the body, mainly of the extensor muscles; the elbows and knees are extended and they resist passive flexing; the tail is stiff and straight; the neck and head are retracted. The condition is undoubtedly due to overactivity of the reflex tonic function of the spinal centers, for it disappears when the posterior spinal roots are cut. The reflexes that depend on the tone of the musculature—for example, the knee-jerk and extensor thrust—are very pronounced in such an animal, and, on account of the higher integration present, reflexes appear that are absent in animals having the cerebrospinal axis cut lower down. For example, although such an animal can not feel, yet when a stimulus is applied that in a normal animal would cause pain, the vocal apparatus may be excited so that a sound or cry of pain is produced. The rigidity does not affect the respiratory muscles. After such an operation, however, normal respiration is much more likely to be maintained if the section is in front of the anterior corpora quadrigemina than behind it.

Removal of the Cerebral Hemispheres.—This furnishes us with what is known as a decerebrate preparation—that is, one in which the animal retains everything from the basal ganglia downward. The operation produces a condition which varies according to the habits of the animal. Thus, in such *fish* as the Elasmobranchs, which depend for their impressions

very largely on the sense of smell, we find that decerebration causes the animal to become completely immobile. It can not seek food because the sense of smell, upon which it ordinarily solely depends, has been destroyed. In a bony fish, on the other hand, decerebration causes very little difference in the behavior of the animal, provided the thalami and optic lobes have been left intact. It continually swims about and is able to distinguish edible from nonedible material.

In the *frog* the result depends very largely upon whether the optic thalami have been simultaneously removed. Even when these structures have been removed along with the cerebrum, the animal at first appears very little different from the normal frog. It springs away when touched, it climbs up an inclined plane, and when thrown in water it swims. It is, however, quite incapable of producing any spontaneous movement, and is in short nothing more than an extremely complex machine, reacting always in exactly the same way to the same kind of stimulus. When the optic thalami are also intact, spontaneous movements are said to be occasionally observed. Such a frog is said indeed to react on the approach of winter as normal frogs do by preparing itself for hibernation, and with spring, to resume its activity and feed itself by catching insects.

In the *bird,* in which the operation of removing the cerebral hemispheres is a very easy one, the movements after decerebration may be quite complicated, particularly if the optic lobes are intact. Such a bird is more active than usual during daylight, but becomes perfectly still in the dark. It is, however, unable to distinguish friends from enemies, and it shows no fear.

As we ascend further in the animal scale, the operation of decerebration becomes very difficult. Goltz, however, succeeded some years ago in removing practically all of the cerebrum from a *dog* by performing the operation in three stages separated by considerable intervals of time. The animal lived eighteen months after the last operation, and during this time it behaved exactly like an automatic machine. All its reflexes were perfectly normal. It could not distinguish objects, but a bright light caused it to close its eyes. During daytime it walked continuously up and down its cage, whereas at night it would sleep and remain perfectly quiet. When food was placed in the mouth, the dog would masticate and swallow in a perfectly normal fashion, and would reject unpalatable food. While asleep, a very loud sound might awaken it, and when a harmful stimulus was applied to the skin, the animal would snarl and growl and attempt to fight the offending object. There were absolutely no signs of pleasure or of recognition of the person that fed it or of fear.

From these results it is in general clear that the brain stem is able to adjust the motor and the visceral reactions of the animal to changes in the immediate environment, but that no power of spontaneous movement is possible. Although in the higher apes and in man removal of any considerable part of the cerebral cortex is impossible, yet we may infer, from the results which have just been considered, that in the higher animals more and more of the action becomes shifted to the motor centers of the cerebrum. Reflexes which in the lower animals involve only a spinal or a bulbo-spinal tract, also involve in the higher forms a cerebral path which is laid down only as the result of experience and education. The newly born infant is able to perform fewer movements than is the case in the lower forms of animal life, but his power of learning new movements is incomparably greater. He inherits less in the way of stereotyped reflexes, but in place of these he possesses innumerable nerve tracts leading through cerebral neurons, through which new reflex responses may be laid down as a result of education.

In connection with these experiments it is interesting to note that in lower animals it can readily be demonstrated that the general influence of the higher on the spinal centers is of an inhibitory nature. Thus, the latent time of the flexion reflex in the decerebrate frog, as judged by the Turck method,* is very much prolonged when a stimulus, such as that produced by a crystal of common salt, is applied to the optic lobes just posterior to the cerebrum. In general, the influence which the cerebrum exercises on the spinal centers is an inhibitory one, whereas that of the cerebellum is augmentatory.

*Turck's method consists in measuring with a metronome the time that elapses between dipping the foot into weak acid solution and the reflex flexion of the leg.

CHAPTER XCV

CEREBRAL LOCALIZATION

Of much greater practical importance than the experiments in which the entire cerebrum is removed, as described in the last chapter, are those in which various parts of it are destroyed or stimulated. From the results conclusions may be drawn regarding the important subject of *cerebral localization.* The effects produced by removal or stimulation of different parts of the cerebral cortex vary considerably, some parts of the cortex being set apart for the control of the motor mechanism of the body, others for the reception and interpretation of afferent stimuli, while others, and these by far the most extensive, are concerned in the correlation or association of the sensory and motor centers. It may be stated in general that: (1) The precentral region of the cerebrum contains the centers of higher thought. (2) The ascending frontal convolution immediately in front of the precentral sulcus contains the chief motor centers, a center being distinguishable for each muscular grouping of the body. (3) The postcentral convolution has to do with the centers for the immediate reception of sensory stimuli, the so-called sensory centers. (4) A large area occupying most of the parietal lobe and part of the occipital is undoubtedly associational in its function, since from it no response can be obtained by stimulation, etc. (5) Behind this, in the occipital lobe, there is a center having to do with the reception of visual impulses. (6) In the upper convolution of the temporo-sphenoidal lobe, is a similar center for hearing.

These centers have been differentiated from one another by anatomical, experimental and clinical research. At present we shall confine ourselves to the *experimental results.* These are obtained by ablation and stimulation, and in considering the results it will be convenient to divide the centers into motor, sensory, and nonreactive.

ABLATION OF THE MOTOR CENTERS

Removal of the cortex from the area which controls the movements of a definite part of the body—say, the arm—will be found to produce an immediate and profound muscular paralysis. The animal does not use the paralyzed extremity for any purpose whatsoever, and yet the mus-

cles do not undergo any more atrophy than can be accounted for by disuse. The extremity does not suffer from any of the nutritional disturbances which we saw supervene upon destruction of the motor center in the cord; and likewise local reflex actions elicited by stimulation of the local receptors are perfectly normal. A pinprick, for instance, causes the usual flexion reflex.

After some weeks the limb begins to recover and can be used in volitional movement. Recovery rapidly progresses until, in the case of the higher apes, it becomes almost complete in a little over four months. It occurs earlier in the lower animals. When a center is destroyed on the cerebral cortex in the case of man, only partial recovery takes place. So that in general we may say that the higher the animal in the animal scale, the less complete will be recovery from the paralysis produced by cerebral ablation.

Regarding *the nature of the recovery,* several possibilities exist: either the nerve centers become regenerated in the destroyed area, or the corresponding area of the opposite hemisphere or some other part of the same hemisphere or the basal ganglia assume the function. Evidence has been furnished by Sherrington and Graham Brown[12] tending to show that the last of these is the most likely cause for the recovery. Thus, it was found, in working on the arm centers on the brain of the chimpanzee, that after complete recovery of the paralysis produced by removal of the center on one side, stimulation of the area that had been removed caused no movements, indicating that no regeneration had occurred, and that removal of the corresponding center of the opposite hemisphere, although followed by paralysis of the arm to which it corresponded, still did not cause any paralysis of the limb which had recovered from the previous operation. To see whether some other part of the gray cortex might have assumed the lost function, the postcentral convolution was removed two months after the removal of the arm centers. Although a temporary weakness of both arms resulted, the voluntary movements were soon as good as before. These results are of course exactly what we should expect from the experiment on the dog, already described in which the cerebral cortex had been entirely removed, and the conclusion that we must draw is that the basal ganglia assume the function of the lost cerebral cortex.

STIMULATION OF THE MOTOR CENTERS

To investigate the effects of stimulation, it is found that the stimulus is best applied by the electrical method, one pointed electrode, called the stimulating, being applied to the area under investigation, and the other,

called the indifferent electrode and consisting of a flat plate, being placed on some other part of the body, such as the skin of the back. This unipolar method gives much finer results than when the ordinary bipolar electrodes are employed.

Before we describe the results which have been obtained by the use of this method, a question arises which it may be well to consider briefly; namely, how do we know that the electric current is really stimulating the center present in the gray matter of the cortex, and not the numerous nerve fibers that constitute the white matter of the brain and along which, between the two electrodes, it is plain some of the electric current must pass? The evidence that we are really stimulating centers is as follows: (1) The latent period for a response produced by stimulating the centers is much longer than that which follows upon direct stimulation of the white matter. (2) Under deep narcosis, as that produced by chloral or morphine, the effect of stimulation of the gray matter is greatly delayed and altered in type; whereas stimulation of the white matter gives the usual response. (3) A weaker current suffices to stimulate the gray matter than that required for the exposed white matter.

In order to demonstrate the movements which follow stimulation of the cerebral cortex, it is necessary, as will be inferred from the preceding remarks, that the animal be not too deeply anesthetized. Furthermore, it is necessary to be very careful in adjusting the strength of stimulation employed, for the results vary considerably accordingly. When the stimulus is of the proper intensity, the movements are located in some particular group of muscles,—for example, those of the thumb or of the hand,—whereas, if the stimulation is strong, the movements spread over much larger areas. As a result of feeble or moderate stimulation, it is found that the muscles which move are those of the opposite side of the body, and that the localization is finer the higher the position of the animal in the scale of development. The movements are perfectly coordinate and purposeful in character, and reciprocal innervation is evident.

There is, however, a marked difference in the reactions obtained by stimulation of the motor cortex and those obtained by eliciting spinal reflexes. For example, the movements produced by stimulation of the cortex are much more readily modified by slight variations in the condition of the animal, the blood supply, the degree of narcosis, etc., than are those elicited by stimulation of receptor neurons. A careful study of this difference has been made in recent years by Brown and Sherrington.[13] They observed the behavior of two antagonistic muscles acting on the elbow when the respective cortical centers were stimu-

lated, and found that the latent periods were very variable, the after-effects indefinite, and inhibition more prominent than excitation. Moreover, the inhibition was more or less independent of the simultaneous excitation of the antagonistic muscle, in which respect it therefore differed from the type exhibited in the reciprocal innervation of the spinal reflexes (see page 814). Nor were the results obtained from a given cortical center always the same; thus, if a point giving a certain response was stimulated immediately after previous stimulation, the result was often reversed; if it was inhibition in the first instance, it might be excitation immediately afterward. But if sufficient time was allowed, then the response was always of the same kind.

By comparing the effect of simultaneous stimulation of an afferent spinal root and of a flexion or extension point on the cortex, it was found that the stimulation of the afferent root when a flexion point was being stimulated augmented the flexion, but when an extension point was stimulated, stimulation of the afferent root might change the response to flexion, the exact result depending considerably on the relative strength of the two stimuli. The general conclusion that may be drawn from these results is that the special function of the cortex is to reverse the centers of purely spinal reflexes when such reversal is desirable or necessary. The cortex dominates the spinal reflexes, and in general it may be said that its main effect is inhibitory in nature.

It is particularly by the use of the method of **moderate electrical stimulation** that exact localization has been worked out on the cerebral cortex. As would be expected, this localization is much less defined and definite in the lower as compared with the higher animals. In the higher apes, it has been found that the motor centers are limited to a narrow strip of cortex immediately in front of the Rolandic fissure—the precentral area, as it is called. (Fig. 219.)

From the accompanying figure, it will be noted that the centers are arranged from below upward, in the reverse order to that in which the muscular groups occur in the body; that is to say, the face, neck, etc., are located lowest on the cortex, and the leg highest up. It will further be noted that the extent of the centers for the neck and tongue is very much greater than for the body or leg, that for the arm being intermediate. It is not, therefore, the extent of the muscular tissue that determines the size of the cortical area controlling its movements, but the type or complexity of the movements that the muscles perform. The complex movements of the tongue and the vocal cords evidently require greater cortical representation than do the coarser movements of the large mass of muscular tissue of the trunk. The centers extend somewhat upon the mesial aspect of the brain, but occupy here only a

very small part of the superficial gray matter. They extend also into the fissure of Rolando and the other fissures, and the extent of the excitable area which is thus buried away in the fissures may exceed that on the free surface of the hemispheres.

It will be noted that there are *two centers for the movements of the eyes,* one in the frontal lobe isolated from the motor area, and the other at the tip of the occipital lobe. The former is the motor center for the conjugated movements of the two eyeballs, whereas the latter functionates in association with the so-called visual center, which receives the visual impressions and transmits them to other parts of the

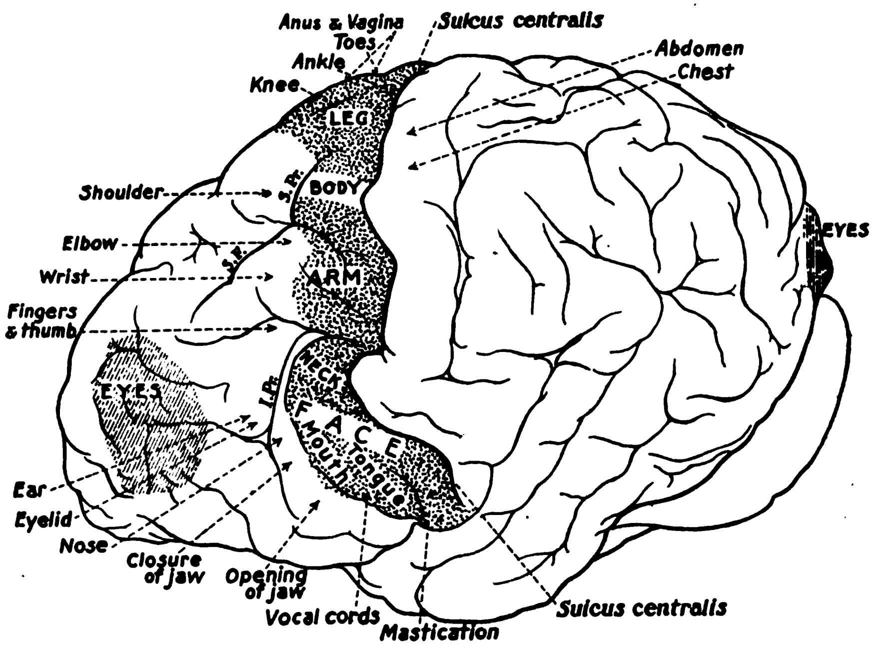


Fig 219.—Outer aspect of the brain of the chimpanzee, showing the position of the centers. Electric stimulation at the parts indicated causes coordinate movements of the corresponding muscle groups. (After Sherrington.)

brain to be interpreted and correlated. Excitation of the center for eye movements in the frontal lobe, say, of the right side, causes conjugate deviation of both eyes to the opposite side, that is, to the left; and it can readily be shown that this movement of the eyeballs is the result of reciprocal innervation of the extraocular muscles (page 814). Even at the risk of repetition we will again describe the fundamental experiment that demonstrates this. When the eyes, as in the above experiment, move to the left, it means that the internal rectus of the right eye and the external rectus of the left are contracting, whereas the external rectus of the former and the internal rectus of the latter

are becoming reciprocally inhibited, the other muscles participating to a slight degree. If all the nerves to the extraocular muscles of the right eye are cut except the sixth, which supplies the external rectus, it will be found that this eye looks permanently toward the right side; that is, an external strabismus is produced. If now the right cortex is stimulated, both eyes will, as before, move to the left, although the right eye will not move farther than the middle line. Its movement as far as this, however, must evidently be due to an active inhibition of the external rectus muscle, for none of the other muscles can act since the nerves are cut.

The experiment of conjugate deviation brings out another point regarding cerebral localization—namely, that the muscles which ordinarily act along with muscles on opposite sides of the body, are innervated from both sides of the cerebral cortex. This applies not only to the movements of the eyes, but to the respiratory and other movements of the neck and trunk. Destruction of the trunk center of the cerebral cortex on one side does not produce any paralysis, while stimulation in this region produces an equal movement on both sides. We may say therefore that bilaterally acting muscles are innervated from both sides of the cerebral cortex.

The movements produced by **stronger stimulation** of the cerebrum do not remain localized, and they persist for some time after the stimulus has been removed. Still further increase in the strength of the stimulus may cause the contraction to spread until it affects all parts of the body, giving rise to a convulsion. There are two types of contraction during this convulsion, the first being a strong tonic contraction, which outlasts the stimulus for some time and then gives way to a series of clonic contractions, occurring at intervals of from six to ten per second, and gradually getting slower as the fit dies away. The convulsion always starts in the muscle group represented by the cortical center that is being stimulated. Thus, if the hand area is the seat of stimulation, the convulsions begin in the muscles of the hand; then they spread to the muscles of the forearm and shoulder on the same side, and then to the face, the trunk, and the leg; and if the stimulus is strong enough, they may spread to the opposite side and thus involve the whole body. This "march" of the convulsion depends upon the overflow of the stimulus on to the various centers of the brain, and the pathways through which it occurs seem to be located in the subcortical centers, for the spread is not prevented by isolating the cortical centers from one another by cuts encircling them, or by division of the corpus callosum. Nevertheless, the centers do in some way become involved in the spread, as is indicated by the fact that the complete excision of one of them

will exclude the corresponding muscular area from participation in the fit.

CLINICAL OBSERVATIONS

The foregoing results obtained by experimental stimulation in animals, are very similar to the symptoms observed in man when the cerebral cortex is stimulated by the pressure on it of a meningeal tumor or a spicule of bone. Such stimulation causes contraction in the corresponding muscular area; the contraction then spreads to neighboring groups of muscles, and may ultimately involve the whole musculature of the body in a convulsive fit, like that produced in animals. This is known as Jacksonian epilepsy, and it is to be distinguished from ordinary epilepsy by the fact that the patient does not become unconscious during the fit. Like ordinary epilepsy, however, the Jacksonian type is usually preceded by a peculiar sensation of numbness or tingling in the area that is to show the first contraction. One of the greatest achievements of modern brain surgery is the cure of Jacksonian epilepsy, by trephining the skull over the affected center and removing the meningeal tumor or spicule of bone which is responsible for the stimulation. To enable the surgeon to locate exactly the position of the irritating body, it is necessary to examine the patient very closely as to the muscular group which is initially affected during the convulsions, and then to examine an outline map of the cerebral hemisphere indicating the position of the various motor and sensory areas as deduced mainly from experiments on the higher monkeys and verified by the experience gained by previous operations. Topographic maps indicating the surface markings corresponding to the various convolutions of the cerebrum must also be used. In such operations the surgeon often has the opportunity of experimentally verifying the position of various centers.

CHAPTER XCVI

CEREBRAL LOCALIZATION (Cont'd)

SENSORY CENTERS

That the motor centers are located in the areas which we have just described does not indicate that the nerve cells of the centers normally dominate the reflex movements which their stimulation elicits. The motor centers, strictly speaking, are the anterior horn cells of the spinal cord; and the so-called motor centers of the cerebral cortex must really represent nothing more than internuncial neurons between the entering and leaving paths concerned in reflex movements. They are only links in the long cerebral chain—way-houses on the reflex cerebral pathway. According to this view we should expect that these centers would be the ultimate recipients of sensation, as well as the distributors of motor impulses; sensorimotor, they have been called. Such, however, is not the case, for Sherrington has shown that the centers most directly concerned in the reception of sensory impulses are not located in front of the Rolandic fissure but immediately behind it *in the ascending parietal or postcentral convolution.* Electrical stimulation in this region does not evoke any immediate response, at least if the stimulus is not too strong. A movement indirectly due to the receipt of a sensation may be elicited by a strong stimulus, just as is the case when the visual center in the occipital lobe is strongly stimulated, producing secondary movements of the eyes.

Histologic, experimental and clinical evidence has been furnished to support this location of the chief sensory center. The clinical evidence was furnished by Harvey Cushing,[14] who induced two patients in whom this part in the brain was exposed to allow him to stimulate it while they were in a conscious state. As the result of the stimulation of the postcentral convolution definite sensory impressions were experienced, consisting of a sensation of numbness or deadness to tactual impressions, but no muscular groups underwent movement unless the precentral convolution was stimulated. During these movements, moreover, no sensations were experienced by the patient except those which accompanied the change in the position of the part that was moved. The sensations which are thus represented on the cortex are those of touch discrimina-

tion and those relating to the position and movements of the muscles. Pain and temperature sensations do not seem to have cortical representation.

There is of course a close association between sensory and motor centers, as is illustrated in the experiment described elsewhere under the head of apesthesia (page 838), in which it will be remembered that the complete section of all the posterior roots of an extremity renders the part as effectively paralyzed for volitional movement as it would have been had the motor roots themselves been cut. Afferent impulses are therefore necessary for the efficient volitional control of the muscular movements.

SENSE CENTERS

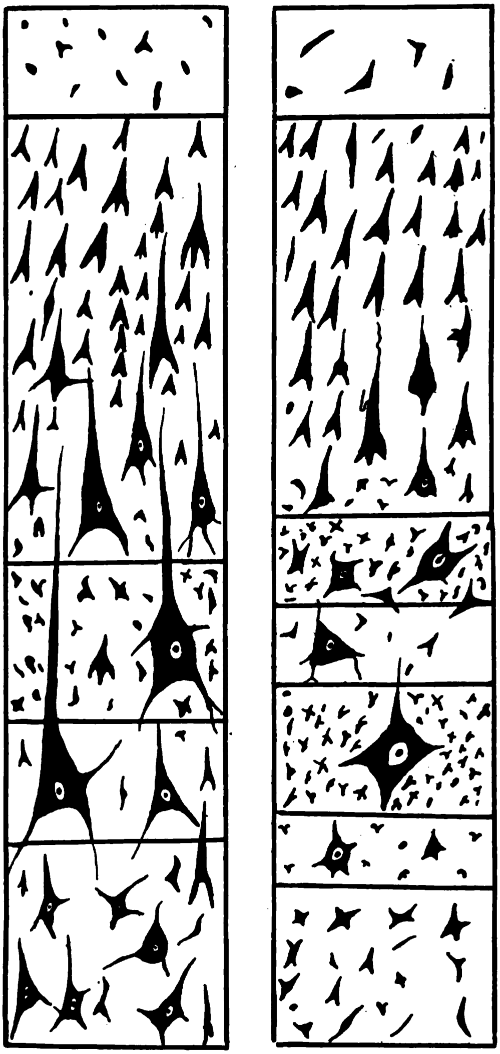
Attempts to locate exactly the position on the cerebral cortex where impressions of the projicient sensations—vision, hearing, etc.—are received are of course more or less difficult because of the fact that the experiments have to be performed on dumb animals. Nevertheless some information can be gleaned from the results of ablation and stimulation of various parts of the cortex, ablation causing, for example, definite evidence either of blindness or of deafness, and stimulation causing movements of the eyes or ears similar to those ordinarily observed when these organs are stimulated in the usual way.

The *auditory center* is located in the back part of the superior temporal convolution. Stimulation of this area in animals causes a pricking up of the ear on the opposite side as if the animal heard a sound. Clinical observation has confirmed this conclusion.

The *visual center* is located in the occipital lobe. It is important to repeat again that there are two centers on the cerebral cortex concerned in vision: the *frontal visual center,* located as we have seen in the frontal lobe, and the so-called *visual center itself,* located in the occipital lobe. Stimulation of the frontal visual center produces a prompter movement of the eyes than does stimulation of the occipital center, indicating that the frontal center has the immediate control of the muscular movements, whereas the occipital lobe is probably concerned in the adjustment of the muscular reactions which are necessary in controlling the eye movements, so that the objects may be properly viewed and judgments formed, by the extent of the movements, of its distance, position, etc. The actual response to stimulation of the occipital centers shows that the lobe on one side is connected with the corresponding half of each retina, the fovea centralis being, however, connected with both lobes.

ASSOCIATION AREAS

The brilliant outcome of the researches conducted by the experimental method in enabling us to locate the chief motor and sensory areas of



Motor leg area Visuosensory Visuopsychic

Fig. 220.—Three sections through different parts of the cerebral cortex. For description see content. (Redrawn from Starling.)

the cerebral cortex leaves yet uncharted those vast areas lying between them which do not respond in any definite way to artificial stimulation, and the ablation of which results only in more or less indefinite symp-

toms. In order to discover the exact function of these areas, it has been necessary to employ an entirely different method—that of histologic and embryologic examination. When the patterns of the gray cortex are compared with the habits of the animals, in different groups of animals (phylogenetic study), or even in different individuals of the

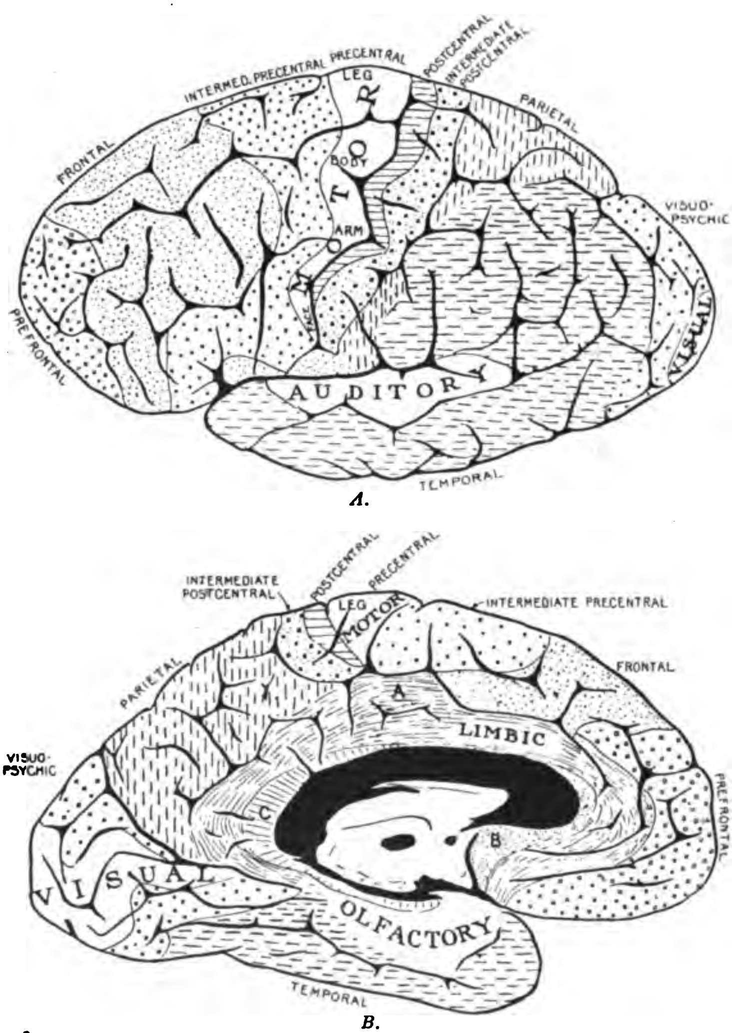


Fig. 221.—The location of the chief motor and sensory areas on the outer (*A*) and mesial (*B*) aspects of the human brain, as determined by the microscopic structure of the cortex. These maps are only approximately accurate, but they indicate in a general way how the cortex is structurally composed. (From Starling after Campbell.)

same group (ontogenetic study), much useful knowledge concerning cerebral localization can also be gained. In the human animal much progress is being made by comparing the structural pattern of the cor-

tex in different parts of the normal brain with that found in the brain of psychopathic individuals whose mental symptoms have been carefully studied before death.*

For the purpose of this work it is necessary to recognize several laminæ or layers of nerve cells and nerve fibers composing the cortex. The most practical division is represented in Fig. 220, and is as follows: (1) a narrow superficial layer of nerve fibers, with a few scattered cells—the outer fiber lamina or molecular layer; (2) a much wider layer of small, medium and large pyramidal cells—the outer or pyramidal cell lamina; (3) a layer of granules—the middle cell lamina; (4) an inner layer of nerve fibers, sometimes containing large solitary cells (Betz cells)—the inner fiber lamina; (5) a layer of polymorphic cells—the inner cell lamina. This five-layer type undergoes structural modifications in the different regions of the cortex, and these modifications possess a distinct functional significance. The only part of the brain in which they can not be recognized is the hippocampus and the pyriform lobe. Based on the relative thickness of these layers maps of the cerebral cortex have been produced. The most important are those of Brodmann and Campbell, of which the latter is reproduced in Fig. 221. Two regions can be very definitely located; namely, the precentral or Betz-cell area, and the visual or visuosensory area of Campbell. Surrounding the visuosensory area is a definite visuopsychic area, and similarly, bordering on the precentral is the so-called intermediate precentral area. At the very front of the frontal lobe is the prefrontal area. Postcentral and intermediate postcentral areas are indicated, but the remainder of the center is undefined.

Reasoning from phylogenetic and ontogenetic data, we can assign to each of these laminæ a functional significance, which according to Bolton is as follows: The outer cell lamina (pyramidal cell lamina) probably constitutes the "higher level" basis for the carrying on of the higher or psychic cerebral functions. It is a prominent feature of the human cortex, the last cell layer to be evolved, and the one which undergoes retrogression most readily. During development it rapidly attains maturity in the visuosensory area, next in the visuopsychic, and only later in the prefrontal region. In the visuopsychic area it is practically of the same depth as in the visuosensory, whereas in the prefrontal region it develops according to the mental capacity of the animal. In patients exhibiting symptoms of dementia this layer of cells is distinctly deficient. These facts indicate that the outer or pyramidal cell lamina "subserves the psychic or associational functions of the cerebrum"—(Bolton).

*An excellent account of the physiologic basis for such work is given by Bolton in Leonard Hill's *Further Advances in Physiology*. We have made free use of this article in the present review and would strongly recommend its perusal by any who may desire further information.[39].

The middle cell lamina is much hypertrophied in the so-called projection areas of the cerebrum—for example, in the visuosensory area (see Fig. 20), where it is so thick that it is usually described as being divided into two parts by a narrow fiber band (the line of Gennari). Diminution in the layer occurs in the visuosensory area in long-standing cases of atrophy. "It seems therefore primarily to subserve the function of receiving afferent impressions whether these arrive directly from the lower sensory neurons or indirectly through other regions of the cerebrum."

The fifth or inner cell lamina is the first to become differentiated, and it is of extremely constant depth in the adult. It is not much affected in amentia, unless when this is very severe, as in patients who are unable to carry on the ordinary animal functions. In short, "it subserves the lower voluntary and instinctive activities of the animal economy"—(Bolton).

Taking the results as a whole, it appears that the region of the cortex behind the Rolandic fissure consists of sensory areas and association areas which may be immediately connected with them (visuopsychic and intermediate postcentral) or more removed (in parietal lobe). The portion in front of the Rolandic fissure, on the other hand, contains the efferent areas, of which the precentral may be regarded as of lowest grade. The motor discharges from it are conditioned upon impulses coming partly from the adjacent intermediate precentral area, in which again are elaborated those received from the sensory areas, and partly from those coming from the prefrontal region, which is the most complex zone of association. This last is indeed the supreme dominating area. It coordinates or integrates the activities of the other association areas and may be considered as the seat of the intellect. The evidence for this high evolution of the prefrontal area is very strong. It is the last portion of the cortex to be evolved and the first to undergo retrogression. In idiots and imbeciles the thickness of the pyramidal cell layer in this region is directly proportional to the mental power, and in dementia degrees of retrogression occur that vary directly with the existing grade of dementia. In normal brains this layer is the very one which varies in depth in different individuals. Along with its high development in the brain of man, as compared with that of other animals, goes hand in hand a great increase in the other association areas. Thought is the product of integration between these various association areas, and articulate and written language the outward manifestation of the process. It is owing to the relatively great extent and complexity and constant development of the prefrontal lobe that man so far excels even the highest apes in his intellectual activity, and it is owing to the relative functional development of this lobe that individuals differ from one another in their mental powers.

CHAPTER XCVII

CONDITIONED AND UNCONDITIONED REFLEXES

In studying the reflexes in the spinal animal, we have seen that the effect of a given stimulus or of different stimuli is predictable with absolute certainty. There is a fatality in the responses. When the higher centers are included in the reflex arc, the reflexes become modified or inhibited by events occurring in other parts of the central nervous system and the results come to be more and more unpredictable. The reflex comes to be a conditioned reflex (Pavlov). Studies of the circumstances affecting these conditioned reflexes really constitute a study of the function of the higher centers in the brain. Since such experiments must be performed on the lower animals, we are limited in the investigation to motor responses, for we have no way whatever of studying the subjective sensations produced. The motor phenomena by which the animal may express its sensations can be interpreted by us only in terms of psychologic ideas that in large part are derived from our own experiences. Obviously the conclusions that can be drawn are subject to very great sources of error, and it must be considered as one of the greatest advances of modern physiology that Pavlov and others should have succeeded in evolving methods by which we may arrive at conclusions regarding the nature of certain of the integrations that occur in the higher centers of the nervous system preceding the motor manifestations by which the animal expresses its sensations.

The methods employed for the study of these higher integrations of the central nervous system all depend on the reactions of the animal that are associated with the taking of food. When the food is actually placed in the mouth, it excites a secretion of saliva, whatever the circumstances may be. This is an unconditioned reflex. Suppose, however, that every time food is given a particular sound is made; after some time it will be found that the occurrence of the sound alone is sufficient to cause a secretion of saliva. In other words, a conditioned reflex has been formed. Similarly, sight or smell or any other type of sensation may be made the excitant for the conditioned reflex. The secretion now becomes psychic instead of merely physiologic. To quote Bayliss: "Any phenomenon of the outer world for which the animal in question possesses appropriate receptors can be drawn into temporary

association with salivary secretion, so that it becomes an exciter of secretion if only it has been frequently presented at the same time with the unconditioned reflex stimulus, food in the mouth."

Work along lines similar to that devised by Pavlov has more recently been undertaken by students of animal behavior, who have utilized the acquired habits of an animal in searching for its food in order to study the influence of conditioning circumstances on its procedure. The advantage of this method depends mainly on the fact that it can be applied to all groups of animals. In carrying out such an observation, the animal is placed in one compartment of a cage, from which it is then released to a second compartment, the end of which is divided into two passageways, one leading to food, the other leading to some compartment in which the animal is punished for its mistake as by receiving an electric shock. Objects such as colored lights are placed in the different passageways, and the animal by repeated trial comes ultimately to learn which particular colored light signifies the passage along which he will receive food. A reflex has therefore become established conditioned on the particular colored light.

On account of the unavailability of his publications, it is impossible at present to give any complete account of Pavlov's discoveries. A few facts, however, are of such importance that it is necessary for us to state them here as far as we know them. (See Bayliss, *Physiology*.) Two mechanisms seem to be concerned in the conditioned reflexes: (1) that of temporary association, and (2) that of analysis. Temporary association is well illustrated in the above experiment in which the secretion of saliva is induced by a sound. Temporary association of the sound with the secretion of the saliva may readily be inhibited by all kinds of external phenomena; thus, if the dog's attention becomes diverted while the conditioned reflex is being stimulated, the response does not occur. In a dog that had been trained to secrete saliva to the sound of a particular metronome beat, inhibition occurred one day because, just as the dog was being presented with the food, the laboratory servant made a noise outside of the building which diverted the animal's attention. The conditioned reflex may also be interfered with by internal inhibition, which is illustrated by experiments in which, after a dog has been trained to respond to a given conditional reflex, several occasions follow when food is not given to the animal after the particular sensation to which it has been trained to respond. The condition—for example, a sound—loses its effect. This is internal inhibition, but it is a temporary condition since the reflex returns of itself after a period of rest.

These experiments illustrate what is meant by the formation of temporary associations occurring in conditioned reflexes, but in order that

CHAPTER XCIX

FUNCTIONS OF THE CEREBELLUM

In our discussion of reflex action we have so far considered only those receptors coming from the exterior of the body, although we have recognized that a considerable number of the afferent nerve roots contain fibers coming from receptors situated in the muscles, the tendons and the joints, and called proprioceptors because they respond not to changes in the environment but to alterations in the body itself. We have seen that the proprioceptors consist structurally of muscle spindles and of the nerve endings in the tendons and ligaments and synovial membranes. They are receptors that are attuned to respond to differences in tension caused either by bulging of the muscles or by stretching of the fibers of tendons and ligaments.

The impulses are transmitted in the spinal cord, either by the posterior columns or by the lateral cerebellar tracts. Those traveling by the posterior columns are sent mainly to the cerebral cortex of the opposite side, whereas those in the cerebellar tracts enter the cerebellum by the inferior peduncles of the same side. The cerebral impulses connect with neurons, which transmit the impulse back again to the cerebellum of the opposite side, so that ultimately the cerebellar cortex is connected with the spinal cord of the same side either directly or indirectly through the cerebral cortex. These anatomic facts indicate in a general way that we may expect the function of the cerebellum to be that of the chief nerve center concerned in the integration of the proprioceptive impulses originated by the condition of contraction or relaxation of the different groups of muscles in the body, and by the amount of tension existing in the various tendons, ligaments and other membranes surrounding the joints.

Experimental investigation has justified these expectations. The removal of the entire cerebellum—an operation which has usually been performed on birds, particularly pigeons, because of the ease with which it can be done in these animals—leads immediately to a condition in which muscular activity is entirely uncontrolled. A pigeon after this operation flies about in an incoordinate way, turning summersaults, dashing itself against the walls of its chamber, and ultimately after constant futile movements, exhausting itself. If one cerebellar lobe is removed, the body when at rest is curved toward the side of the lesion, and the movements of the

animal cause it to fall in that direction. A similar experiment with dogs yields like results, but the operation is of course considerably more difficult. In man, a destructive tumor of the cerebellum produces a condition known as "cerebellar ataxy," in which the patient moves his limbs in a very incoordinate fashion; he staggers, is uncertain in his gait, and behaves in general very like a drunken man.

Although these immediate effects of cerebellar extirpation indicate

Fig. 222.—Footprints after destruction of the cerebellum in a dog: *a*, before the operation; *b*, four days after; *c*, five days after; *d*, a month after; *e*, two months after. (From Luciani.)

clearly that this organ has to do with the control of muscular movements, yet the results are probably not primarily dependent on the ablation, but rather on the conditions of irritation which are set up as a result of the operation, and which probably affect the cerebellar peduncles. At least, such is the view that Luciani, one of the greatest investigators in this field, has adopted because of the fact that, if the animal is kept alive for sufficient time so that the symptoms of irritation disappear, they become

replaced by those of an entirely different nature. The pigeon may reacquire the power of flying straight, or—and this is particularly important—the dog may reacquire the power of apparently normal progression, although, if its muscular movements are carefully examined by physiologic methods, it will be found that at least three changes have developed as a late result of the extirpation; namely, a weakness of contraction, a tremor during the contraction, and a want of tone when at rest. These conditions have been called *asthenia, atonia* and *astasia,* respectively. On superficial examination it may often be difficult to make out these three conditions, but they can readily be observed in animals in which the cerebellar extirpation has been performed on one side, so that the abnormal may be compared with the normal side. In a dog that has had one cerebellar hemisphere removed some time previously, the muscles on the corresponding side are so much weaker than those on the opposite side that the animal, in order to retain his equilibrium, has to prop himself up either by leaning against whatever object may be convenient, or by extending his legs so as to increase his base of support. In other words, he constantly tends to fall to the side of the lesion, but tries to prevent this either by increased muscular effort or by taking advantage of artificial support. The effect which this weakening has on his gait can be very clearly demonstrated by comparing the footprints produced by the normal with those of the abnormal side, these footprints being obtained by making the animal trot along a piece of glazed paper blackened with a carbon deposit as in taking tracings (Fig. 222).

Localization of Function in the Cerebellum

Although these facts in themselves would tend to indicate a certain degree of *localization of function* in the cerebellum, or at least that certain parts of the cerebellar cortex have to do with certain groups of muscles, yet for many years it was considered that the cerebellum did not show in any marked degree the same kind of localization that we find in the cerebral cortex. One cause for the backward state of our knowledge concerning cerebellar localization is that, unlike the cerebrum, its cortex is practically inexcitable. In recent years, however, on account partly of anatomic and partly of experimental and clinical work a high degree of localization has been found to exist in the cerebellum. From the anatomic point of view it has been found that in certain groups of animals, such as the ungulata, the postero-medial lobule of the cerebellum is very large, whereas the lobuli ansiformes are small. In another group, the carnivora, the opposite obtains, the lobuli ansiformes being greatly developed and the postero-medial lobule quite small.

By studying these **developmental differences** in relationship to the activities of the muscular system, Bolk suggested that movement of those regions of the body which are affected by muscle groups on both sides—for example, the head, neck or trunk—would be represented on the cerebellar cortex by an unpaired center—that is, a center occupying a middle position—and that this would be capable of exercising an influence equally upon the muscles of both sides. Movements of the limbs would require an entirely different type of coordination, since they are not accustomed to act together, unless for certain movements, as walking. Based on these theoretic considerations Bolk found a definite correspondence to exist between the variations in the development of certain cerebellar lobules and the functional importance of certain muscle groups, and the general conclusions deducible from his and

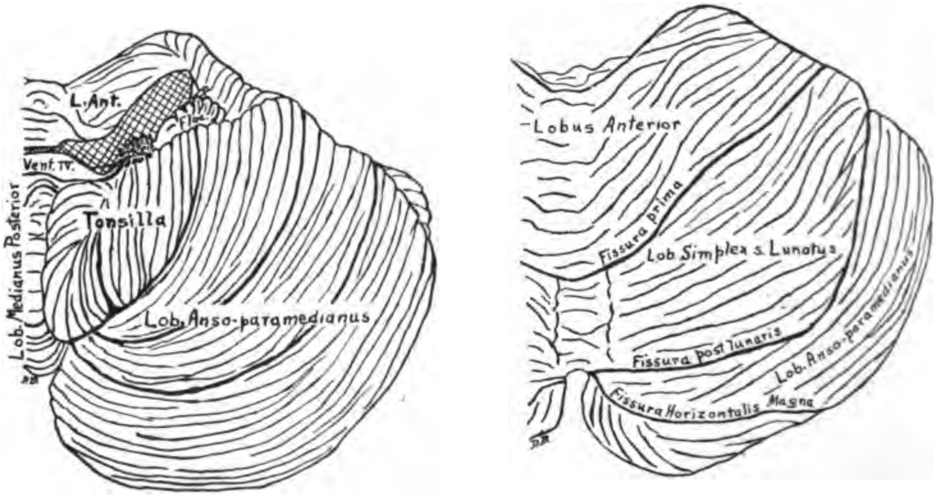


Fig. 223.—Diagrams to represent respectively a ventral view of the left half and a dorsal view of the right half of the human cerebellum illustrating the scheme of subdivision according to Bolk. (From photographs of specimens from the Anatomical Museum, Western Reserve Medical School.) (From Davidson Black.)

correlated work may be summed up as follows (cf. Davidson Black[16]): The lobus anterior cerebelli (see Fig. 223) contains the centers for the coordination of the muscle groups of the head (eyes, tongue, muscles of mastication, muscles of expression), and of the larynx and pharynx. The lobus simplex contains centers for the coordination of the muscles of the neck. The lobulus medianus posterior contains the unpaired centers for the synergic movements required by the right and left extremities for the purposes of progression. On the other hand paired centers for the extremities—those centers that have to do with the independent movements of each limb of the same side of the body—are located in the lobuli ansiformes et paramediani (crus primum and crus secundum). The centers for the rest of

the trunk and tail region are located in the remainder of the cerebellum. These conclusions are the basis of the accompanying map of the cerebellum.

Basing his work on these anatomic conclusions, Van Rijnberk has studied the effect of **circumscribed extirpation** of certain lobules of the cerebellum on the muscular control of the different parts of the body, with the following results. Total or partial extirpation of the lobulus simplex produces side to side oscillations of the head, indicating the removal of the influences of the cerebellum that control the movements of the muscles of the neck. Complete extirpation of the crus primum of the lobuli ansiformes causes as an immediate—irritative—effect dynamic disturbances of the fore limb of the same side, replaced later by a condition of atonia, which makes the limb hang limp, and of asthenia, which makes it feeble in its movement when it is excited to contract. Extirpation of the crus secundum has a similar

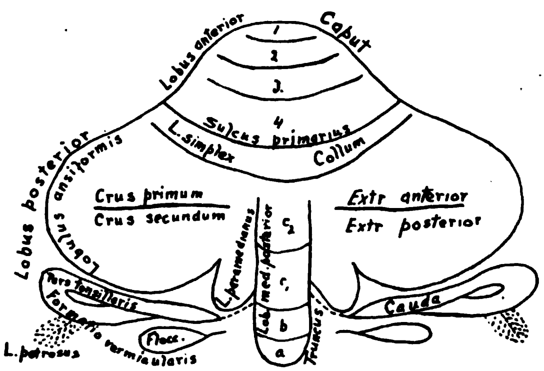


Fig. 224.—Schema of the parts of the mammalian cerebellum spread out in one plane. (After Bolk by Van Rijnberk from Luiciani. Op. cit.) On the right side of the figure the relation of the different lobules to the functional development of the musculature is indicated according to the theory of Bolk noted in the text. (From Davidson Black.)

influence on the muscles of the hind limb of the corresponding side. Extirpation of both crura of the lobulus ansiformis causes marked asthenia and atonia in both fore and hind limb on the same side as the lesion. A characteristic disturbance in walking develops as a late effect of this extirpation. It has been termed the "hen's gait." Extirpation of the lobulus paramedianus causes rotation on the longitudinal axis of the body, with pleurothotonus to the operated side. (Fig. 224.)

Just as in the case of cerebral localization, so in cerebellar we find that within each of the largest centers a more particular localization can be made out; thus, in each of the centers for the upper and lower extremities, there is a definite arrangement of subsidiary centers for the direction of the activities of antagonistic muscle groups concerned in the movements of particular joints. It must be remembered, however, that in all these cases no real paralysis is produced by extirpation, but only a want of coordina-

tion on account of the fact that the sthenic, tonic and static impulses required for muscular harmony are not properly elaborated.

After some time, as in the case of complete cerebellar extirpation, the symptoms gradually disappear, but they can be obtained more or less characteristically in practically all animals, at least in all those that have been investigated, including dogs and monkeys.

It will be of interest to consider for a moment the possible causes for the ultimate disappearance of the symptoms of cerebellar extirpation. These are either: (1) an organic compensation by the uninjured parts of the cerebellum, or (2) a functional compensation by the voluntary centers of the cerebrum. Although the former of these methods of compensation may sometimes develop after partial destruction of the cerebellar cortex, it can not of course explain the recovery which we have seen to occur after the entire cerebellum has been removed. The most important compensation no doubt is effected by the cerebrum, as the following observation clearly indicates. If half of the cerebellum of a dog is destroyed, and the animal kept alive until the symptoms of cerebellar extirpation have entirely disappeared, it will then be found, if the cerebral center on the opposite side is removed, that the symptoms return in their original severity. After this second operation the powers of standing in the erect position and of walking are permanently lost.

CLINICAL OBSERVATIONS

Application of these laboratory results has been recently made in the clinic, the most important contribution having come from the clinic of Barany, who for his work was awarded the Nobel prize. In cases of abscess, cysts, or regional agenesia, it is now possible to determine the exact site of the lesion in the cerebellum. To effect this localization, it has been necessary to work out certain clinical tests. The most important of these is called the *index test*. This is described by Davidson Black as follows: "The patient's eyes being closed, he is asked to execute a simple movement in a given direction with one of his extremities. For example, the forearm being firmly supported, the patient's index finger is extended and brought into contact with that of the observer; the patient is then required to move his finger vertically downward and then to return it to its previous position. The test is repeated a number of times, both in the vertical and in the horizonal direction, and if any tendency toward deviation from the plane of movement be present, its direction is noted. By slight modification of the foregoing procedure it is possible to test each of the limb segments in all positions of rotation, pronation or supination."

The index test is applied (1) without previously producing nystagmus

and (2) after producing artificial nystagmus. The artificial nystagmus is produced by spinning the patient two or three times, and consists of constant lateral movements of the eyeballs, quick in the direction in which the artificial movement took place, and slow in the opposite direction.

In a normal subject, previous to spinning the index test shows no deviation, but after the production of artificial nystagmus a deviation is noted in the direction corresponding to the slow jerk of the nystagmus (reaction

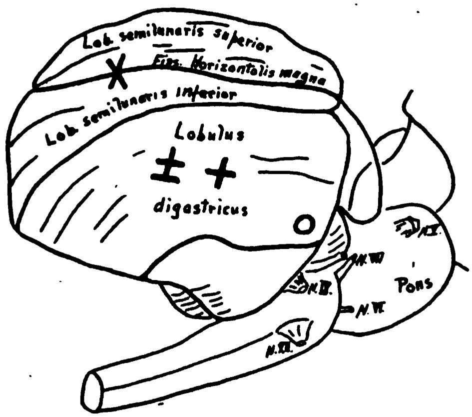


Fig. 225.

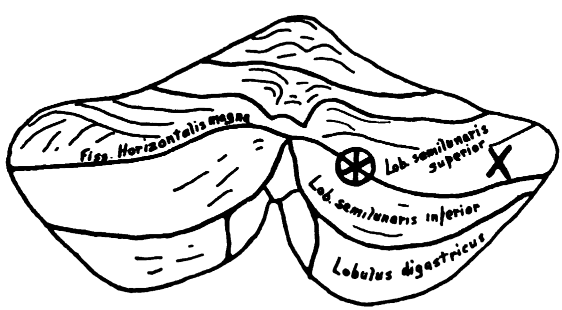


Fig. 226.

Figs. 225 and 226 represent respectively the inferolateral and the posterior aspect of the human cerebellum indicating certain cerebellar localizations according to Barany. (After Barany, from André-Thomas et Durupt. Op. cit.) N. VII, Nervus facialis; N. IX, Nervus Glossopharyngeus; N. XII, Nervus hypoglossus.

The signs in the above diagram indicate the exact localization of the centers for the tonus of the musculature concerned in some of the movements of the right arm and leg, ⊗ marks the center for downward movements of the arm; ×, for abduction of the arm; O, adduction of the hand; + adduction of the arm; ±, adduction of the hip. N. V. indicates Nervus trigeminus; N. VI, Nervus abducens; N. VII, Nervus facialis; N. IX, Nervus glossopharyngeus; N. XII, Nervus hypoglossus. (From Davidson Black.)

deviation). When a cerebellar lesion exists, the index test performed on a patient without nystagmus sometimes causes a spontaneous deviation in the segment of the body corresponding to the position of the lesion on the cerebellar cortex, but more frequently, if it is applied after the production

of artificial nystagmus, the reaction deviation in that segment will fail to be obtained. The exact site of the cerebellar lesion is diagnosed partly from the nature and direction of the deviation which is produced and partly from the segment of the limb in which it occurs, the explanation for the disturbances being that interference with the cerebellar control of one muscle group causes the antagonistic muscular groups to perform their movements in an exaggerated manner, so that the segment moves too much in their direction.

Barany's conclusions so far may be summarized as follows:

(1) The centers for the extremities are located on the cortex of the hemispheres in the semilunar (superior and inferior) and digastric lobules (see Fig. 225). The representation is uncrossed or homolateral, thus contrasting with cerebral localization, in which it is crossed or heterolateral.

(2) Within each of these chief centers there is a further localization, which however does not refer to anatomic groups of muscles but rather to the functional performances of the different segments of the limb. Thus, within the arm centers there are subsidiary centers concerned in the movements of the limb in the various planes in rotation, in pronation and in supination. It is a functional rather than an anatomic localization.

(3) When a center concerned in the movements of the limb in a certain direction, e. g., to the right, is suddenly destroyed, a spontaneous deviation is produced in the opposite direction (to the left).

CHAPTER C

THE CEREBELLUM AND THE SEMICIRCULAR CANALS; FUNCTIONAL TESTS

The cerebellum serves as the great nerve center to which are transmitted, through the various proprioceptors, the impulses which, as it were, inform it as to the exact degree of muscular effort required to maintain the animal in its various postures. It is, as Sherrington puts it, the head ganglion of the proprioceptive system. Such impulses from the muscles, tendons, etc., could not, however, supply information regarding the exact position of the body in space. For this purpose special receptors connected with the eighth nerve are provided in the *semicircular canals*. These, it will be remembered, are three in number on either side, each canal consisting of a semicircular bone tube attached to the vestibule of the internal ear; and they are arranged so that they lie at right angles to one another in the three planes of space. The three canals on either side are thus disposed so as to form an arrangement like a V-shaped armchair with the back inwards. This arrangement causes the posterior vertical canal of one side to be in the same plane as the superior vertical canal of the opposite side, the external canals being in the horizontal plane on both sides. The arangement will be plain from the diagram (Fig. 227).

Within the osseous canals are suspended membranous tubes, which do not fill the canals. The canals, etc., contain fluid, but are not completely filled. The osseous as well as the membranous canals are dilated at one end to form the ampulla, and it is here that the vestibular division of the eighth nerve ends in a structure called the "crista acustica," consisting of hair cells supported by sustentacular cells. The nerve terminates in fine arborizations between the hair cells. In the saccule and utricle are structures similar to the crista, called the maculæ acusticæ. These structures are receptors specialized for the purpose of responding to changes in the position of the head, and therefore of the body in general. When the head moves in a certain plane of space, the fluid in the membranous canals and in the utricle and saccule on account of inertia undergoes a certain movement, which acts on the hairs of the hair cells and thus sets up a stimulus. According to the degree of the stimulation in the various ampullæ, which again will be dependent upon the plane or planes in which the movement of the head has occurred, impulses are transmitted

through the vestibular nerve, and these impulses ultimately reach the cerebellum.

The experimental evidence for these conclusions regarding the functions of the semicircular canals is very strong. Thus, after destruction of all the canals—an operation which is particularly easy in the pigeon—the animal behaves very much the same as after cerebellar destruction. After some months these disorders disappear, because the cerebellum learns to control the movements of the body from other proprioceptive impulses, particularly those of sight. If such a recovered animal is bandaged, the symptoms return in all severity. This compensation is furthermore an educative process, for it does not occur when the cerebral centers as well

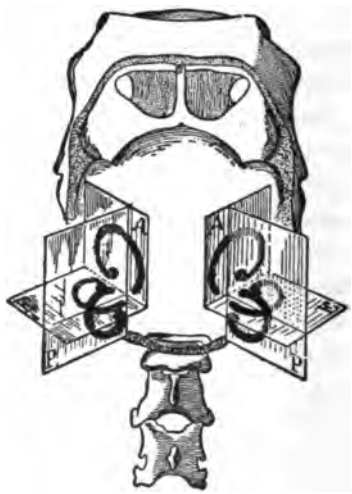


Fig. 227.—The semicircular canals of the ear, showing their arrangement in the three planes of space. (From Howell's *Physiology*.)

as the semicircular canals are removed, and it can be abolished in a recovered animal by removal of the cerebral cortex.

Many observations of great interest have been made concerning these labyrinthine sensations by Pike and others, but we can not discuss them further here. One point of interest, however, is that forced movements in definite planes are induced by removal of a canal. Removal of the horizontal canals, for example, causes continued nodding movements of the head in the plane of the injured canals. An experiment of great significance was performed by Ewald to show the effect of causing a movement of the fluid in one of the canals. For this purpose a bony canal was opened at two places by a dental drill. Through the hole farther from the ampulla, amalgam was introduced so as to block the backward movement of fluid, and into that nearer the ampulla a fine tube was inserted connected with a rubber bulb. By manipulating the bulb, the membranous

canal could be compressed and currents set up in the endolymph. It was found that these currents always caused the animal to move its head and eyes in the plane of the canal that was being stimulated and in the direction of the current of endolymph.

Finally, visual impressions supply much of the information the cerebellum requires, the close association of the movements of the eyeballs with cerebellar and labyrinthine disturbances being well recognized. The nystagmus already described in connection with Barnay's tests is dependent upon this association. The symptoms and sensations of giddiness or nausea produced by rotation of the body, or by unusual movements such as those of a boat, are no doubt due to the irregular and unusual variety of labyrinthine sensations which they excite.

In a word, then, the function of the cerebellum is to receive proprioceptive impulses from the body along with labyrinthine and visual impressions and to integrate and develop from them impulses which, by being transmitted to the cerebral and other nerve centers that dominate muscular movements, so coordinate the nerve discharges from them that, when muscular movement occurs, it does so in relationship to the previous position of the animal and in the most efficient way to attain the object for which the movement was made. The cerebellum is the head nucleus of the proprioceptive system.

THE ASSOCIATION BETWEEN THE EYE MOVEMENTS AND THE SEMICIRCULAR CANALS

The close association between the eye movements and the semicircular canals is indicated by the occurrence of nystagmus when the ear is stimulated either electrically or by means of moderately cold water impinging on the membrana tympani. The latter method of inducing nystagmus is styled the caloric, and is employed in the examination of candidates for the aviation service. Its value over the tests of nystagmus after rotating the body and the index test already described depends on the fact that it enables us to test each vestibular apparatus separately.

Water at 68° F. is allowed to run through a stop nozzle into the external auditory canal, free of wax, from an irrigation bottle placed about 3 feet above the head, which is meanwhile tilted at an angle of 30° forward. In about 40 seconds a rotary nystagmus with the direction of the jerk to the side opposite to the douched ear should be evident, or dizziness complained of. The reaction test is then applied and immediately afterward the head is inclined at an angle of 60° backward from the perpendicular, when a horizontal nystagmus to the side opposite to the douched ear should develop. Deviation is again tested.

The procedure is repeated on the other ear. If it takes longer than 90 seconds for the nystagmus to appear, the vestibular apparatus of that side is abnormal. Absence of the reaction deviation after the douching is a certain sign of internal ear disease.

It is undoubtedly essential that these tests should be most carefully applied to all would-be aviators. They frequently reveal lesions of the vestibular apparatus or the cerebellum in subjects who had thought themselves perfectly normal, and who indeed may have boasted of their powers of equilibrium because they imagined that freedom from seasickness or failure to become dizzy in dancing indicated a high development of this function. There can be no doubt that many aviators have gone to their death because of impairment in the ear mechanism. When on "terra firma" the muscular sense and cutaneous sensations often make the vestibular weakness of no consequence, but when deprived of these contributory sensations and dependent on the ear-balance mechanism alone, as in flying, any weakness becomes a serious menace.

CHAPTER CI

THE AUTONOMIC NERVOUS SYSTEM

In discussing the physiology of the central nervous system, we have broken away from the traditional textbook treatment of the subject in that we have left practically untouched any description of the course of the various nerve tracts or of the position of the nerve centers. We have pursued this policy in the belief that the study of these details of structure belongs just as surely to the anatomist as does the structure of other parts of the body, notwithstanding that to trace the course of the nervous pathways he may have to call to his aid the physiologist and clinical neurologist. There is one part of the nervous system, however—namely, the involuntary or autonomic—the physiology of which it is impossible to discuss apart from its anatomy, because this has depended very largely on physiologic methods for its elucidation. Until such methods were emphasized and while anatomy alone was depended upon, little could be learned of the functions and connections of the sympathetic chain and of the various nerve plexus that compose the involuntary nervous system. We shall here review briefly the general anatomic plan of this system as described by Gaskell.[17]

GENERAL PLAN OF CONSTRUCTION

The plan of the involuntary nervous system is much the same as that of the voluntary, the main points of difference being dependent upon the location of the neurons composing the reflex arcs. It will be remembered that there are three of these: the receptor, the internuncial, and the effector neurons (page 782). The receptor neurons have the same position for both systems; namely, the posterior root ganglia (Fig. 228). The internuncial neurons of the involuntary system, like those of the voluntary, have their cells in the spinal cord, where they are represented, in the thoracic region, by the cells of the lateral horn of gray matter; in the sacral region, by a similarly placed collection of cells; and in the bulbar region, mainly by the dorsal nucleus of the vagus. The main cause for the difference between the two systems is dependent on the course of the fibers of the internuncial neurons; in the involuntary system they leave the spinal cord *before* connecting with the effector neuron nerve cells, which are contained in the various ganglia found throughout the body, whereas

in the voluntary, they remain within the spinal cord, and terminate on the effector neurons, which are the anterior horn cells.

The outflow from the spinal eord of involuntary internuncial fibers, which we shall hereafter style *connector* fibers, occurs along the anterior spinal roots, but is somewhat irregular in distribution, because it is interrupted in the cervical and lumbar regions, where the nerve plexus for the extremities come in. There are, therefore, three main regions of outflow for the connector fibers—a thoracicolumbar, a bulbar, and a sacral; and the fibers (Fig. 229) do not behave in the same manner in all of them. The fibers of the thoracicolumbar region form the so-called *sympathetic system,* and run by the corresponding white rami communicantes either immediately to the ganglia of the sympathetic

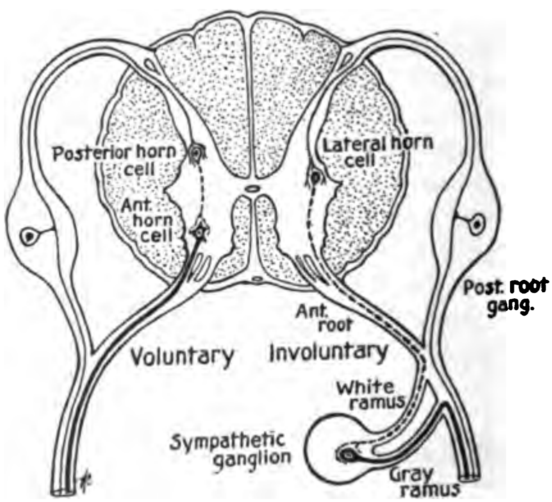


Fig. 228.—Diagram to illustrate the different arrangements of the internuncial neurons of the voluntary and involuntary nervous systems. In both systems the afferent fiber terminates (by collaterals) around a cell of the gray matter of the cord. In the voluntary system this cell is situated in the posterior horn, and its axon travels to an anterior horn cell. In the involuntary system, on the other hand, it is located in the lateral horn, and its axon leaves the cord by the anterior root and travels by the white ramus into a sympathetic ganglion, where it connects with a nerve cell, whose axon forms the postganglionic fiber. (From Gaskell.)

chain, or by the splanchnic nerves to the abdominal ganglia. In the ganglia are situated the cells of the effector neurons. The fibers of the sacral region connect with effector neurons, forming the *pelvic ganglionic group* (pelvic nerves, nervi erigentes); and those of the bulbar outflow with effector neurons located either peripherally or in the ganglia of the *vagus* and the seventh, ninth, and eleventh cranial nerves. In the midbrain there is a fourth group of involuntary connector fibers running to effector neurons found in the ciliary ganglia.

The anterior roots of many of the spinal nerves are therefore not

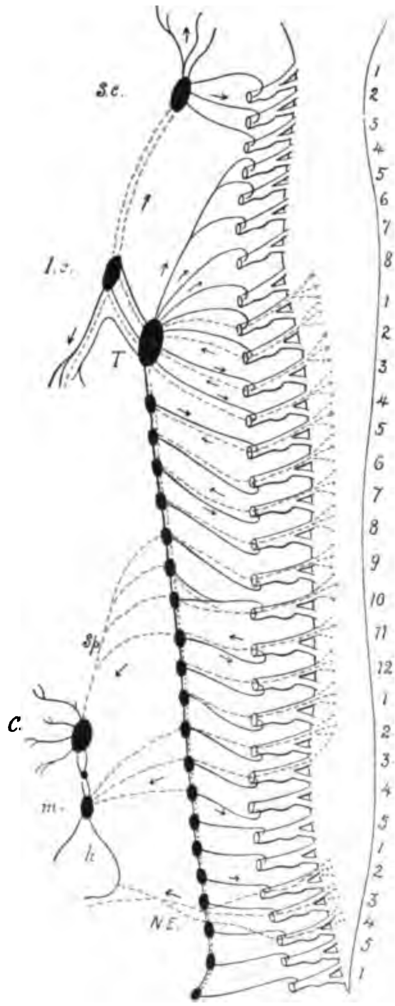


Fig. 229.—Diagram of the sympathetic nervous system to be used along with Fig. 233. The preganglionic fibers are in red, and the postganglionic in black. *S.c.*, superior cervical ganglion; *I.c.*, inferior cervical ganglion; *T*, stellate ganglion; *S.p.*, great splanchnic nerve; *C*, ganglia of solar plexus; *m.*, inferior mesenteric ganglia; *h.*, hypogastric nerves; *N.E.*, nervus erigens. The arrows indicate the direction of nerve conduction. The numerals indicate the spinal nerves. (From Howell.)

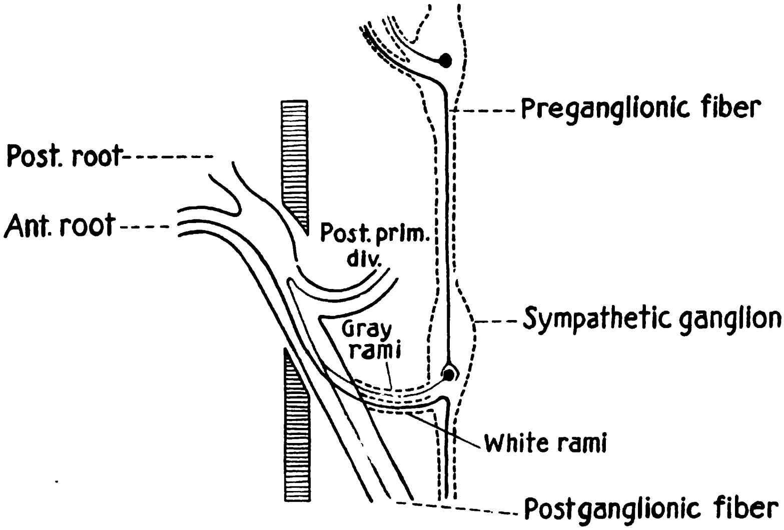


Fig. 231.—Diagram (after Langley) to show the manner in which a preganglionic fiber, emanating from the spinal nerve by the white ramus communicans, connects in a ganglion of the sympathetic chain with a nerve cell (red), the axon of which then proceeds as the postganglionic fiber (red) by way of the gray ramus communicans back to the spinal nerve, along which it travels to the periphery. It will be observed that the preganglionic fiber does not form its synapsis in the first ganglion it encounters.

stellatum for the fore limb, and the lower lumbar and upper sacral ganglia for the hind limb. When they are stimulated, the muscular fibers surrounding the sweat glands contract and squeeze out the sweat.

SITUATION OF BLOOD VESSELS	SITUATION OF MOTOR GANGLION CELLS	ROOTS CONTAINING CONNECTOR NERVES
Head and neck.	Superior cervical ganglion.	1, 2, 3, 4, 5, thoracic; 2, 3, 4, give maximum effect.
Heart.	Ganglion stellatum and inferior cervical ganglion.	1, 2, 3, 4, 5, thoracic; 2, 3, give maximum.
Anterior extremity.	Ganglion stellatum.	4, 5, 6, 7, 8, 9, thoracic, and 10 slightly.
Posterior extremity.	6th lumbar, 7th lumbar, and 1st sacral ganglion.	11, 12, 13, thoracic; 1, 2, lumbar and 3 slightly.
Kidney.	Renal ganglion.	4, 5, 6, 7, 8, 9, 10, 11, 12, 13, thoracic; 1, 2, 3, 4, lumbar.
Spleen.	Semilunar ganglion.	3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, thoracic; 1, 2, 3, lumbar.
Abdominal viscera.	Superior mesenteric ganglion and semilunar ganglion.	6, 7, 8, 9, 10, 11, 13, thoracic; 1, 2, 3, lumbar.
Pelvic viscera.	Inferior mesenteric ganglion.	1, 2, 3, 4, lumbar.

(Gaskell)

The ganglia for the *pilomotor fibers* are more widespread (extending from the fourth thoracic to the coccygeal ganglia); but the connector fibers are again strictly confined to the thoracicolumbar region. Stimulation of these fibers causes movement of the hairs, or on hairless animals, the condition called "goose skin."

The Motor Nerves of the Muscles Surrounding the Segmental Duct.—It will be observed that the connector fibers to the abdominal and pelvic viscera are collected into two special nerve trunks, the greater and the lesser (or lumbar) splanchnics. The collateral ganglia (semilunar and superior and inferior mesenterics) with which these connect, have nothing to do with the segmental nerves, but their nerve cells send fibers (postganglionic) which supply the various viscera not only with vasomotor fibers but also with the "sympathetic" fibers, which we have seen exercise such an important control over their glandular and muscular functions.

All of the fibers contained in the lumbar splanchnics do not, however, have their cell stations in the inferior mesenteric ganglia, but run through it and proceed in the hypogastric nerves to find their effector cells on the musculature of the various structures that are developed from the Müllerian and Wolffian ducts—i. e., of the ureters, uterus, Fal-

lopian tubes and vas deferens. Many of the fibers of the hypogastric nerves are therefore those of involuntary internuncial neurons.

The *ileocolic and internal anal sphincter muscles* of the intestines and *internal vesical sphincter* receive their nerve supply from effector neurons situated in the superior and inferior mesenteric ganglia, the internuncial fibers arising from the thoracicolumbar region. It is possible that the other sphincters of the intestinal canal—viz., the cardiac and pyloric sphincters of the stomach—are similarly innervated. (Fig. 232.)

Great aid in working out these nerve connections is received by studying the effect of epinephrine, which acts specifically on those tissues that are supplied by the sympathetic nervous system.* Epinephrine has no effect on tissues innervated by the bulbar or sacral outflows, and it develops its action peripherally, being indeed more potent on a denervated organ even after all its nerves have been allowed to degenerate. Advantage of this action of epinephrine has been taken in the investigation of doubtful cases of sympathetic innervation, such as in the cerebral, coronary, and pulmonary blood vessels. The outcome of these investigations has been discussed elsewhere.

THE BULBOSACRAL OUTFLOW OR THE PARASYMPATHETIC SYSTEM

From the medulla oblongata arise involuntary connector neurons, which are carried mainly by the vagus nerves but partly by the seventh, ninth and eleventh cranial nerves to effector nerve cells situated peripherally on the structures to which the nerves run (Fig. 233). These include in a general way the muscles and glands of the alimentary canal and its derivatives as far as the end of the small intestine. In the small intestine itself the cells of these motor neurons are those of Auerbach's plexus found between the two muscular coats. In the diverticula, which include the lungs and the gall bladder, the nerve cells to which the vagus fibers run are also situated peripherally.

The *sacral outflow* occurs through the second and third sacral nerves, the fibers joining to form a single nerve (the pelvic nerve or nervus erigens) on each side. This passes directly to the bladder, where it connects with a plexus, often called the hypogastric, which extends over the bladder and neighboring portion of the rectum. The branches run to connect either with the nerve cells of the ganglia of the plexus itself, or with nerve cells situated on the walls of the large intestine and bladder. The pelvic nerve makes its connections with the periphery in the

*Its action is always the same as that which is produced by stimulation of the sympathetic nerve supply, whether this effect is one of stimulation or inhibition.

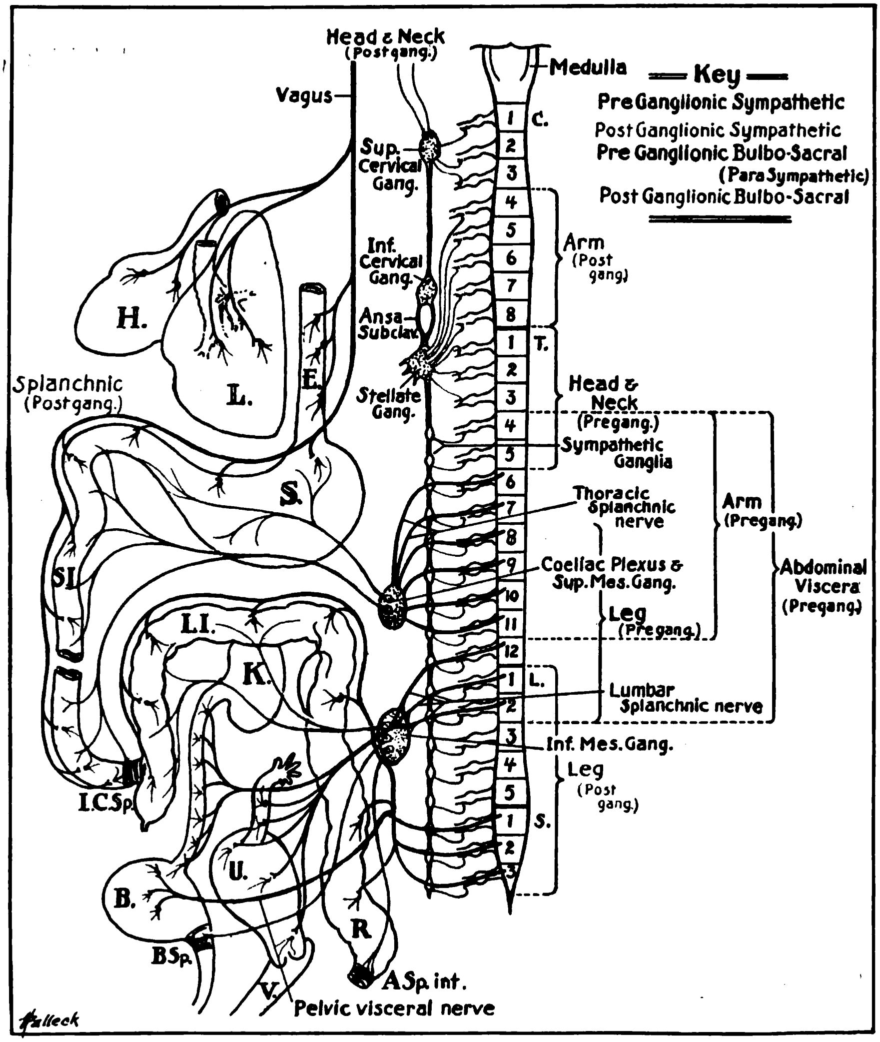


Fig. 232.—Diagram showing the main parts of the autonomic nervous system to be used along with Fig. 233. For the sake of clarity several of the preganglionic fibers of the sympathetic autonomics are omitted, but the position of their egress from the cord is indicated in the side notes. The diagram shows clearly the distribution of the bulbosacral autonomic system by way of the vagus and the first, second and third sacral nerves.

same manner as the vagus. Taken together these two nerves supply the musculature of the gastrointestinal tract, including the cloaca, the vagus as far as the end of the small intestine, and the pelvic nerve from this point on. It must of course be remembered that certain muscles—namely, the sphincters of the small and large intestine—receive their nerve supply from the sympathetic (page 882). Just as structures innervated by the sympathetic are peculiarly susceptible to the action of epinephrine, it has been discovered that those innervated by the bulbosacral system are very susceptible to the action of acetyl-choline, which is present in ergot. They are not acted on by epinephrine, nor are the structures upon which this acts affected by acetyl-choline.

AXON REFLEXES

At this place it is convenient to consider for a moment the phenomenon which has already been referred to as an axon reflex. It was discovered that when one of the hypogastric nerves was cut and the central end stimulated there was a reflex contraction of the bladder and the internal anal sphincter, along with vasoconstriction in the region of the rectum and that this occurred, even after disconnecting the inferior mesenteric ganglion from the spinal cord by cutting the lumbar splanchnic nerves. Injection of nicotine immediately abolished the response. It looked as if reflex action was possible through the ganglion; which would justify the name "sympathetic" originally given to the involuntary nervous system in the belief that the ganglia were centers for local reflex actions. Further investigation showed, however, that this reflex is not similar to those occurring in the voluntary system, but is dependent upon the presence of a collateral on internuncial fibers that run through the inferior mesenteric ganglia to nerve cells situated peripherally on the walls of the bladder and rectum. The collaterals terminate by synapsis around nerve cells in the ganglion, the axons of which, as we have seen, run to the bladder, the rectal blood vessels, and the internal sphincter ani. The evidence for this explanation depends on the observation that the axon reflex is no longer possible after the lumbar splanchnics have been cut and time allowed for their fibers to become completely degenerated.

Similar reflexes depending on collaterals have been found in the lateral chain, and there can be little doubt that they are of frequent occurrence throughout the whole involuntary nervous system, just as they are, within the spinal cord, in the voluntary. It is because of these collaterals and the fact that nerve fibers transmit impulses in both directions that a stimulus transmitted through one or a limited number of connector

neurons may excite a broad field of effectors and cause a widespread effect.

FUNCTIONS OF AUTONOMIC NERVES

The functions of the autonomic nerve fibers have been discussed in connection with the structures which they supply, and we shall require in this place only to review them in a general way.

Two opposed effects may be obtained: *stimulatory* (augmentory) and *inhibitory;* and these may be produced through one nerve by its being stimulatory for one set of muscle fibers and inhibitory for another set in the same viscus. The branches running from the inferior mesenteric ganglion to the colon, for example, are augmentory (constrictor) for the blood vessels and inhibitory for the muscular walls of the colon.

The greatest interest centers on the inhibitory impulses. They are best known in connection with the vagus nerve to the heart, the sympathetic to the small intestine, and the hypogastric to the musculature of the bladder. It is interesting to compare the nature of inhibition in the involuntary and voluntary systems. In the latter, it will be remembered, inhibition can occur only through the internuncial neurons and the effector nerve cell, stimulation of the effector nerve fiber never having any other than an augmentor effect. It is quite otherwise in the involuntary nervous system, for stimulation of the effector nerve fiber, after complete destruction of the effector nerve cell, is still followed by a typical inhibition. This, it will be remembered, may be demonstrated on the frog heart by applying electric stimulation to the white crescentic line after paralyzing the effector nerve cells by nicotine. The same may also be shown in the case of the chorda tympani, where stimulation of the post-ganglionic fibers in the hilus of the gland causes dilatation of the blood vessels after paralysis of the ganglion by nicotine, vasodilatation being of course a phenomenon of inhibition.

It is a difficult matter to designate precisely which fibers in any part of the involuntary nervous system are inhibitory and which augmentory. Indeed, as mentioned above, one fiber may perform both functions. In cases where the existence of inhibitory fibers is doubtful, great aid is afforded by the use of ergotoxine, an alkaloid of ergot, which possesses the remarkable property of specifically paralyzing the augmentor nerves of the sympathetic system (but not of the parasympathetic); that is, the same fibers as are stimulated by epinephrine. When a particular structure is supplied with augmentory and inhibitory fibers by a combined sympathetic nerve, electric stimulation or the application of epinephrine usually gives only augmentory effects; after the injection of

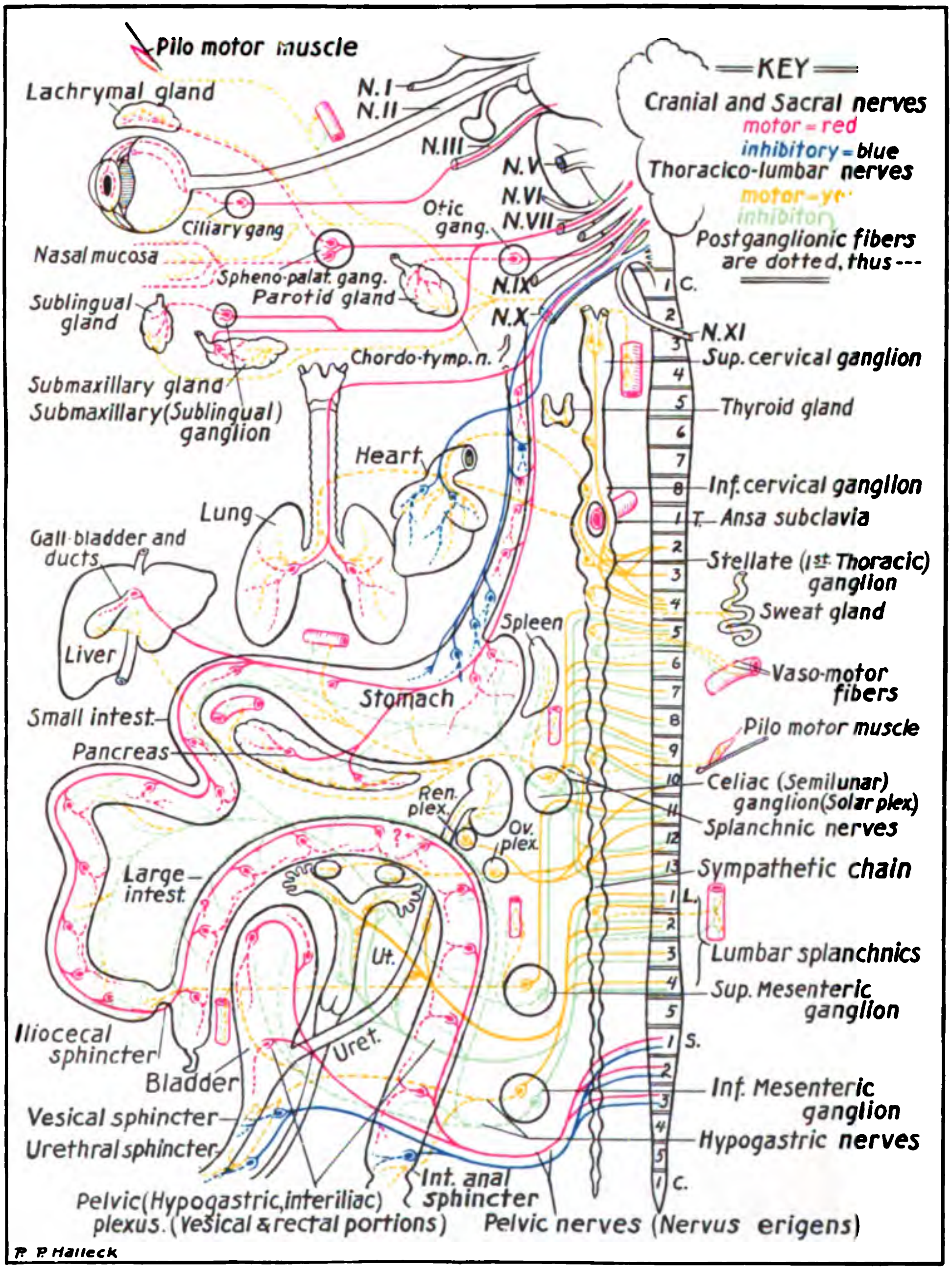


Fig. 233.—Schematic representation of the involuntary nervous system. (From Jackson.)

ergotoxine, however, a reversed effect is observed; namely, inhibition instead of augmentation. By taking advantage of this fact, Dale has been able to demonstrate in the hypogastric nerves inhibitory fibers to the uterus, and Elliott has demonstrated the inhibitory action of epinephrine on the muscles of the ureter in the dog. Inhibitory fibers have also been discovered by these methods in the great splanchnic nerves, in the nerve roots supplying the kidney, and in the cervical sympathetic supplying the blood vessels of the mucous membrane of the mouth, etc.; that is, in nerve trunks which previously were believed to contain only augmentory fibers. The accompanying diagram from Jackson will give an idea of the currently accepted views concerning the distribution of augmentory and inhibitory fibers. (Fig. 233.)

THE AFFERENT FIBERS OF THE AUTONOMIC SYSTEM

It has long been known that the exposed viscera are remarkably insensitive. This experience is in accord with the observation that the supply of afferent fibers to the viscera is relatively very small. In the hypogastric and probably in the great splanchnic nerve, Langley computes that only about one-tenth of the medullated fibers are afferent. At the two ends of the alimentary canal, where cooperative reflexes between the somatic musculature and the viscera are of importance, a greater number of afferent fibers are found in the autonomic nerves; for example, in the pelvic nerve about one-third of the fibers are afferent, and, as we have frequently seen, the vagi contain large numbers of them coming from the lungs, stomach, and no doubt from other viscera.

The afferent visceral fibers, as above stated, arise like those of the voluntary system, from the ganglion cells of the posterior roots. They travel in company with the connector fibers through the white ramus communicans, so that the stimulation of the central end of one of these may cause reflex rise in blood pressure and other movements.

It is found that, after opening the abdominal cavity under local anesthesia, cutting and suturing of the viscera may be continued without causing any pain. When the viscera are inflamed, however, and under certain conditions of stimulation, such as the distention of the bile ducts with biliary calculi, or the violent contraction of the intestines, excruciating pain may be evoked. This pain is frequently not localized in the viscera, but is referred to certain parts of the surface of the body, and it has been shown by Mackenzie and by Head that it is referred to the area of skin which is supplied with sensory nerves by the same segment as that to which the afferent autonomic fibers of the particular viscus run. It has further been shown that vascular disease may cause sensi-

tivity of the corresponding cutaneous areas, so that clinical methods are available for localizing the site of the disease by studying the exact position and extent of the referred pain or skin tenderness.

NERVOUS SYSTEM REFERENCES

(Monographs and Original Papers)

[1]Parker, G. H.: Proc. Am. Philos. Soc., 1911, i, 217-225.
[2]Head, H., and Rivers, W. H. R.: Brain, 1908, xxxi, 323-450.
[3]Meek, W. J.: Am. Jour. Physiol., 1911, xxviii, 356-360.
[4]Bruce, A. Ninian: Arch. f. exper. Path. u. Pharmakol., 1910, lxiii, 426-433.
[4a]Sherrington, C. S.: Numerous papers on reciprocal innervation of antagonistic muscles, Proc. Roy. Soc., Vol. B, 66; also in Jour. Physiol., xxii, xxxiv, xxxviii, xliii, and Quart. Jour. Exper. Physiol., ii.
[5]Holmes, Gordon: Brit. Med. Jour., 1915, ii, Nov. 27, Dec. 4 and 11.
[6]Pike, F. H.: Am. Jour. Physiol., 1909, xxiv, 124-152.
[7]Jolly, W. A.: Quart. Jour. Exper. Physiol., 1910, iv, 67-87.
[8]Lombard, W. P.: Jour. Physiol., 1889, x, 122-148.
[9]Collier, J.: Lancet, April 1, 1916, 711.
[10]Ranson, S. W., and von Hess, C. L.: Am. Jour. Physiol., 1915, xxxviii, 128.
[11]Head, H., and Thompson: Brain, 1906, xxix, 537.
[12]Sherrington, C. S., and Brown, T. Graham: Jour. Physiol., 1913, xlvi, Proc. Physiol. Soc., p. xxii.
[13]Brown, T. Graham, and Sherrington, C. S.: Proc. Roy. Soc., 1912, 85, B, 250-277.
[14]Cushing, Harvey: Proc. Am. Physiol. Soc., Am. Jour. Physiol., 1909.
[15]Luciani, L.: Kleinhirn, Ergebnisse der Physiol., 1904, i.
[16]Black, Davidson: Cerebellar Localization in the Light of Recent Research, Jour. Lab. and Clin. Med., 1916, i, 467.
[17]Gaskell, W. H.: The Involuntary Nervous System, Monographs on Physiology, ed. by E. H. Starling, Longmans, Green & Co., 1916.

Other Monographs not Specifically Referred to in the Context

[18]Sherrington, C. S.: (1) The Integrative Action of the Nervous System, Silliman Lectures, Yale University. Scribner's Sons, New York. (2) Shäfer's Textbook of Physiology, II. Young J. Pentland, London, 1899.
[19]Bolton, J. S.: Recent Researches on Cortical Localization and on The Function of the Cerebrum in *Further Advances in Physiology,* ed. by Leonard Hill, London, E. Arnold, 1909.

INDEX

B

C

D

H

I

N

R

S